Firm Dynamics and Selection in the Labour Market: Evidence from Colombia

Michael R Koelle*

[This version: May 2014]

Work in progress.

Abstract

The chances of being offered an attractive job in wage employment affect the performance of microenterprises. To illustrate how, I embed a simple model of investment under uncertainty in a labour market search framework. Under frictions, initial misallocations between self-employment and wage-employment persist and comparative advantage shapes expectations about transitions out of self-employment, which in turn influence firm performance. In order to test these predictions, I develop a semi-structural model of transitions that accounts for both correlated unobservable heterogeneity, and search frictions. Using a recent panel survey of households in urban Colombia, I find evidence for this mechanism: a 10% increase in the likelihood of an income increase in wage-employment reduces firm growth by 4%. The results from the dynamic search model document that frictions are substantial (and probably larger than in developed countries). With this approach, my paper seeks to build a bridge between the studies of microenterprises as firms, and the study of competition and ‘segmentation’ in labour markets in developing countries.

Keywords: Occupational choice, search frictions, entrepreneurship,

JEL Codes: J24, J64, L26, O17.

*Address: Department of Economics, University of Oxford, Manor Road Building, Oxford OX1 3UQ, e-mail: michael.koelle@economics.ox.ac.uk. Thanks to Marcel Fafchamps, Simon Quinn, and Margaret Stevens, for advice; Paolo Falco, Clément Imbert, Russell Toth, for fruitful discussions; Leander Heldring for comments on an earlier draft; seminar audiences at CSAE 2014, Oxford, and CoDE II, Mannheim; and Adriana Sabogal at Fedesarrollo, for permitting me access to the data they collected. All errors are my own.
1 Introduction

Developing countries are characterised by a large share of self-employment. Many individuals are self-employed or own a small firm with a few employees. These microenterprises are of a multifaceted nature: on the one hand they resemble firms, on the other hand they constitute a means to provide income for their owner.

Different literatures (discussed shortly) have looked separately at these facets of self-employment and microenterprises. One literature has empirically studied microenterprises as firms. In recent years, it has used experimental methods and tends to find substantial heterogeneity in alleviating potential constraints of these firms. Often, effects on outcomes such as profits or growth are basically zero for a large proportion of microenterprises. These findings suggest that further dimensions of firm heterogeneity exist which have not been identified yet.

In contrast, the theoretical and empirical literature on labour markets in developing countries has devoted much attention to whether those markets are competitive or not. In particular, it has asked whether there are sectors in the labour market that offer an earnings premium. This is most likely to be the case in wage-employment. If there is such a premium, individuals in other sectors should be expected to seek entry into the premium sector. This literature has not explicitly dealt with the consequences of non-competitive markets on the behaviour of market participants, especially the self-employed.

I propose to study the implications of these facets of microenterprises jointly: What does it mean for the dynamics of microenterprises if some of the self-employed seek entry into wage-employment? What kind of labour market structure would give rise to such entry-seeking behaviour? And is it possible to identify variation in the aspirations to enter the wage sector that permits an empirical analysis?

1.1 Contribution

Using a recent panel survey of households in urban Colombia, I find that employment growth of microenterprises is inversely related to the chances of

---

1 Throughout, I use the terms ‘self-employed’ and ‘firm owners’, and respectively ‘microenterprises’, ‘firms’ or ‘businesses’, interchangeably.
firm owners to receive an attractive job offer in the wage sector. To explain the mechanism behind these findings, I embed a simple investment model for firm owners into a wider labour market where individuals can earn income through self-employment and wage-employment; those two categories are called ‘sectors’. Transitions between sectors are affected by search frictions. The predictions from such a model are supported by the data: self-employed with a higher probability of receiving an attractive offer in wage-employment are more likely to make this transition, but less likely to grow their firms. By imposing the structure of this search model I am able to quantify these frictions. Furthermore, I find that both frictions and selection on comparative advantage determine the allocation between self- and wage-employment.

Conceptually, this paper offers a novel framework to think about interconnections in labour markets of developing economies. I unite the investment decision of an individual, currently self-employed, with the decision about which form of employment to seek in the future. In the presence of frictions, some individuals are misallocated between sectors. Increasing the probability of switching to wage-employment reduces incentives to make irreversible investments.

This conceptual innovation contributes to various strands of literature and provides a way to explicitly link them. First, I add to the literature on the structure of labour markets in developing economies, especially to the strand of literature which has argued in favour of the presence of structural frictions that lead to ‘dual’ (Lewis 1954), ‘segmented’ (Harris and Todaro 1954) or ‘staged’ (Fields 1975) labour markets. There has been a renewed interest in empirical tests of the implications of these models that exploits some newly available panel data from developing countries, as evidenced by the work of Bosch and Maloney (2010), Günther and Launov (2012) and Haywood and Falco (2012).

I formalise market frictions using a model based on Mortensen (1970) that has long become standard in labour economics. This model decomposes labour market transitions into a component determined by search frictions, and a component of self-selection on comparative advantage. Combining the search model with a two-sector Roy (1954) model of correlated earnings potentials, I derive estimable equations and can empirically distinguish between the components of transitions.

Second, I add to the literature on the determinants of growth of micro
and small enterprises. The neoclassical theory on entrepreneurship (Lucas 1978, Kihlstrom and Laffont 1979) focuses on the selection according to comparative advantage, given a set of constraints (Jovanovic 1982, Hopenhayn 1989, Evans and Jovanovic 1989). Recent experimental studies are able to shed light on the existence and importance of such constraints for microenterprises, with a corresponding focus on capital (de Mel et al. 2008, 2012; Fafchamps et al. 2011), entrepreneurial skills (Karlan et al. 2012), or both (Karlan and Valdivia 2011).

I argue in this paper that the performance of micro and small firms is affected by the labour market in which the self-employed find themselves. This market gives them the choice of alternative employment opportunities, but is also characterised by search frictions that pose additional constraints. In that sense, I propose a way to think about how economic forces that one would not at first associate with a ‘firm’ might still matter for firm outcomes. In this manner I try to explain patterns of firm performance with insights about market frictions from the first literature.

Third, I contribute to the line of studies that examine degrees of heterogeneity in objective and ambition among the self-employed. The fact that small business owners have very diverse goals and backgrounds has already been noted in the early descriptive studies by Liedholm and Mead (1999). Recently, it has been separately acknowledged in anthropology (Coletto 2010), in business studies (Légrand et al. 2012) and in economics (de Mel et al. 2010).

Empirically, this paper provides evidence that supports the conceptual framework, using a panel survey of urban households in Colombia from 2007 until 2010. It also contributes to empirical studies of the Colombian labour market. The two central hypotheses implied by the framework are tested and cannot be rejected. Self-employed are more likely to switch to wage-employment if the chances of earnings gains are higher. Moreover, if such individuals stay in self-employment, they are less likely to grow their firms. The findings are robust to including aggregate shocks, to allowing and correcting for selection on unobservables, and to other alternative explanations. When I estimate the dynamic search model, I find substantial frictions. Although a large proportion of self-employed could improve their income in wage-employment, only a minority receive a job offer every year. I also find evidence of (at least partial) selection according to comparative advantage.
With my analysis, I add to the literature on the Colombian labour market (Magnac 1991, Mondragón-Velez and Peña 2011, Cuesta and Bohórquez 2011). Unlike Magnac, I am able to reject the hypothesis of competitive markets with a dynamic compared to a static structural model, and studying a quite different time period. I find, similarly to Mondragón-Velez and Peña, a heterogeneity in the nature of self-employment in Colombia, though on a continuous scale.

Finally, I provide a simple estimable structural model of labour market outcomes under search frictions, and perform the estimation using maximum likelihood. This model simultaneously allows for search frictions and self-selection based on comparative advantage. Its estimation assesses the relative importance of frictions and choice in shaping labour market allocation. This contributes to the literature of structural estimation of labour market search model (Caluc et al. 2006, Jolivet et al. 2006).

1.2 Outline

The paper is structured as follows. The theoretical framework in section 2 proposes a choice model for the self-employed, with the novel feature of embedding it into the labour market search model. The empirical methodology is described in detail in section 3, including derivation of a dynamic search model. The data are described in section 4. Section 5 presents the main results obtained with reduced-form approaches, and a number of alternative specifications and robustness checks. Section 6 discusses the dynamic estimation results. Finally, section 7 concludes. An appendix contains some of the lengthier derivations, proofs, and an overview of variable definitions.

2 Conceptual Framework

This section presents a conceptual framework that nests a model of investment behaviour of a microentrepreneur within a standard two-sector labour market search model where frictions are present. The model describes an individual who decides whether to be self-employed or wage employed, and how much capital to invest. Each period, she may be able to draw a wage offer, and accepts if being wage employed at that wage makes her better off than being self-employed. However, she has to invest before knowing the realisation of the stochastic elements that determine employment; hence she maximises on expectations about future employment states. Investment has positive returns in self-employment and zero return in wage employment.
Therefore, the model predicts that the more likely wage employment, the lower investment will be.

Consider an individual who earns income from being either wage-employed or self-employed, who has a single asset $k$ for which there is no financial market for borrowing or saving, and who derives utility $u()$ from consumption.\(^2\) Individuals face an infinite horizon and discount with rate $\beta$.

The timing of the model is as following. There are two states individuals find themselves in at the beginning of each period $t$. They are either self-employed, or wage-employed at wage $w$, and they have a level of assets $k$. If they are self-employed, they earn income $\pi(k)$ from a production technology with decreasing returns to scale in the only input, capital. They must decide how much capital $k_{t+1}$ to carry over. The remainder is consumed. Only after they invest, they learn whether they will receive a wage offer (this happens with probability $\lambda$) from a distribution $F(w)$. If they receive an offer, they can decide whether to take it and become wage-employed.

The set of decisions for the self-employed is not as rich. They just earn income $w_t$, consume, and then randomly (with probability $\delta$) lose their job and become self-employed, or stay on in wage employment. Under these simple assumptions the problem has only a single 'true' state variable vector, made up of $(WE, w)$ and $(SE, k)$ pairs. This ensures existence of an analytic solution, which could not be found otherwise.\(^3\)

\(^2\)It is quite easy to extend this model to allow for savings on financial markets. Unless there is some reason for risk diversification, individuals would then choose a cutoff strategy based on financial and entrepreneurial rates of return, and invest only in one asset at a time. Identification of this model relies on microentrepreneurs whose firms still grow and have not yet reached there steady-state size. This is modeled straight away. Borrowing constraints can arise endogenously if income is not strictly bounded above zero, and there is no default (Acemoglu, 2009). Here, this possibility arises for individuals who are thrown into self-employment, without capital. There cannot recur to a subsistence income.

\(^3\)Simulation of a multi-state model, which allows for richer behaviour on part of the wage-employed, is a topic for further extension of this paper I am currently working on.
The model can be stated as a stochastic discrete programming problem:

\[ V(WE, w) = u(w_t) + \beta \cdot \{ \delta V(SE, k_t) + (1 - \delta) V(WE, w) \} \quad (1) \]

\[ V(SE, k_t) = \max_{k_{t+1}} \left\{ u(\pi(k_t) + k_t - k_{t+1}) + \beta \cdot \left[ \lambda \cdot \max \{ V(WE, w), V(SE, k_{t+1}) \} \right] \right\} \quad (2) \]

\[ V(SE, k_{t+1}) = \max_{k_{t+1}} \left\{ u(\pi(k_t) + k_t - k_{t+1}) + \beta \lambda \cdot \int_{z(k_{t+1})}^{\infty} \left( V(WE, w) - V(SE, k_{t+1}) \right) dF(w) \right\} \quad (3) \]

Since \( V(WE, w) \) is monotonically increasing in \( w \), there is a reservation wage strategy for any \( k \). Since \( V(SE, k) \) is also locally increasing in \( k \), the reservation wage is an increasing function of \( k \) for intermediate levels of \( k \). Assuming \( \pi(k) \) has decreasing returns to scale, \( z \) is a concave function of \( k \) for intermediate levels of \( k \). Define the reservation wage, in function of any \( k_t \), as

\[ V(SE, k_t) = V(WE, z(k_t)) \quad (4) \]

which can be expressed as following in the value function:

\[ V(SE, k_t) = \max_{k_{t+1}} \left\{ u(\pi(k_t) + k_t - k_{t+1}) + \beta \lambda \cdot \int_{z(k_{t+1})}^{\infty} \left( V(WE, w) - V(SE, k_{t+1}) \right) dF(w) + \beta V(SE, k_{t+1}) \right\} \quad (5) \]

I am interested in comparative statics of the parameter \( \mu = \Pr(w > z(k_{t+1})) \) on the firm size choice \( k_{t+1} \) of the self-employed. After some manipulation (which is relegated to Appendix A) the first order condition of the value function with respect to \( k_{t+1} \) obtains as follows:

\[ 0 = -u'(\pi(k_t) + k_t - k_{t+1}) + \frac{\beta \lambda}{1 - \beta(1 - \lambda)} \left[ -u'(z(k_{t+1})) \left( 1 - F(z(k_{t+1})) \right) \right. \]

\[ \left. + \frac{\beta}{1 - \beta} u'(z(k_{t+1})) \frac{\partial z(k_{t+1})}{\partial k_{t+1}} \right] + \frac{\beta}{1 - \beta} u'(z(k_{t+1})) \frac{\partial z(k_{t+1})}{\partial k_{t+1}}. \quad (6) \]

\[ ^4 \text{At very low levels of } k, \text{ there is a poverty trap. The best option for self-employed individuals is to run down their capital and take up the next best wage offer. A higher } \mu \text{ may allow them to do so slower, and wait longer for a job in wage employment that pays relatively better.} \]
This FOC defines an implicit function $F$ in $F(k_{t+1}$ and $\mu$. Its signs can be shown (see Appendix A) to be

$$\frac{\partial F}{\partial k_{t+1}} < 0$$

(7)

$$\frac{\partial F}{\partial \mu} < 0.$$  

(8)

Now, we apply the implicit function theorem to obtain

$$\frac{\partial k_{t+1}}{\partial \mu} = -\frac{F_\mu}{F_{k_{t+1}}} < 0$$

(9)

Investment is therefore lower with individuals who face a higher probability to receive a wage offer which is above their reservation wage. Individuals expect that they will take up such an offer and leave self-employment. Accordingly, they make less investments into their firm’s capital stock.

2.1 Predictions

This subsection summarises the testable hypotheses that the above model predictions

**Hypothesis 1.** The probability to switch into wage-employment is higher when the probability of an acceptable job offer is higher.

This hypothesis is immediately predicted by the labour market search model. The probability of switching is given by $\psi = \lambda \cdot \mu + (1 - \lambda)$ and the comparative static is

$$\frac{\partial \psi}{\partial \Pr(w > z(k))} = \lambda > 0.$$  

(10)

**Hypothesis 2.** Credit-constrained microentrepreneurs with a higher probability to receive an acceptable wage offer make fewer investments into their firm.

This hypothesis is immediately predicted by (9).

**Hypothesis 3.** If the production function is homothetic, constrained entrepreneurs are less likely to add employees to their firm.

This hypothesis is immediately predicted by homotheticity in combination with Hypothesis 2.
3 Empirical Framework

In this section I present the identification strategy and the equations I estimate. The main concern is to express the self-employed’s desire for switching into wage-employment, which is an unobservable variable if entry into wage-employment is rationed. Guided by the theory in Section 2, I present two measures of opportunities in wage employment based on earnings differentials. These measures can be identified from data on incomes, and sectors of employment. I address two potential concerns for identification of wage distributions: omitted variables, and selection on unobservables. I show that omitted variables pose no problem for identification of the conditional moments of the wage distribution (as opposed to identification of causal parameters). Furthermore, I discuss identification under alternative sets of assumptions about the dependence of unobservable wage determinants, and the relevant information set for predictions. I propose a semi-structural dynamic switching model (DSM) for identification under weaker conditions than those required for standard approaches.

3.1 A simplification

In order to test the hypotheses summarised in section 2.1 we need to have an empirical measure of the probability that a self-employed individual finds a given job offer acceptable. The search model implies a cutoff strategy, and this probability is given by $\Pr (w_i > z)$. The threshold $z(k)$, however, is only implicitly and recursively defined by

$$
\frac{u(z(k_t)) - \beta u(z(k_{t+1}))}{1 - \beta} = u(\pi(k_t) + k_t - k_{t+1}) + \frac{\lambda \beta}{1 - \beta(1 - \delta)} \times \int_{z(k_{t+1})}^{\infty} u'(w) \left( 1 - F(w) \right) dw
$$

(11)

There is no analytical solution for $z$. Any solution requires numerical simulation methods, and structural assumptions about the shape of $u()$, and values for $\beta$, $\lambda$ and $\delta$. For the purpose of the reduced-form analysis to follow, I assume the reservation wage $z$ to be some general, increasing function of current income from self-employment:

$$
z(k) = g(\pi_t)
$$

(12)

The simplest functional form of $g()$ is $g(\pi_t) = \pi_t$. This approximation is justified when $k_t$ is close to $k_{t+1}$, and when the influence of future option
values on the optimal stopping strategy is small. This would be the case, for example, when the probability of receiving a job offer $\lambda$ is low; or when there is on-the-job-search also in wage-employment, and the offer probabilities are similar across employment states.\footnote{This is outside the present framework since for simplicity the search model does not consider the possibility of on-the-job search by the wage-employed. Such a model is presented, for example, by Mortensen and Pissarides (1999).}

### 3.2 Expressing Alternative Employment Opportunities

The object of interest that allows testing the hypotheses predicted by the model is

$$\Pr(w_i > \pi_i | \Psi_i)$$

This is the probability that a job offered to a self-employed individual is acceptable by her, conditional on the information set $\Psi_i$. This information set generally contains observable characteristics $x_i$, but potentially also other information about the self-employed individual, for example unobservable components of their income. I discuss different information sets below. In line with the model, I interpret (13) as a measure of outside opportunities of a self-employed individual. By construction, this object embeds assumptions about informational frictions, since it implies that self-employed individuals do not know the exact realisation of their potential, counterfactual wage. Wage dispersion arises as a result.\footnote{It is beyond the scope of this analysis to explain the formation of wages. In this partial equilibrium analysis, I merely take a distribution of wages for given. Neither do I delve into the distinction between the wage offer distribution, and the realised wage distribution. For an exposition of how wage dispersion can be explained by labour market search theory, see the book by Mortensen (2003), with a special emphasis on bargaining theory the survey article by Manning (2011) and with a detailed description of the theories of wage posting the survey by Mortensen and Pissarides (1999).}

I also use a simpler measure of alternative employment opportunities. It indicates whether the expected wage $E(W_i)$ is higher than the current income from self-employment, $\Pi_i$:

$$\mathbf{1}(E(w_i) > \pi_i | \Psi_i)$$

where $\mathbf{1}(\cdot)$ denotes the indicator function. This measure sorts the self-employed into just two categories, based on a relative evaluation of earnings: those who wish to leave self-employment, and those who wish to stay.
I next discuss empirical identification of these measures of outside opportunities, before moving on to discuss identification of the effect of these measures on outcomes.

3.3 Identification of counterfactual wage distributions

The central challenge to identification is that we only ever observe an individual in a single sector at a time. In order to assess the relationship between alternative employment opportunities and outcomes for the self-employed, we need to be able to construct the measures (13) and (14) for all self-employed. It is therefore necessary to predict the unobserved, counterfactual wage, and its distribution, for the self-employed. To this I now turn.

First, I set up the problem. Suppose that log wages of individual \( i \) are determined according to the following process:

\[
\ln W_i = \log w_i = \tilde{x}'_{1i}\beta + \epsilon_i \tag{15}
\]

where \( \tilde{x}_{1i} \) is a \((M \times 1)\) vector of explanatory variables which may include a constant, \( \beta \) is a \((M \times 1)\) vector of parameters, and \( \epsilon_i \) is a stochastic error term. We cannot expect to be able to observe all elements of \( \tilde{x}_{1i} \). I make this explicit by writing the earnings function as

\[
w_i = x'_{1i}\beta_1 + x'_{2i}\beta_2 + \epsilon_i \tag{16}
\]

where \( x_{1i} \) is a \((K \times 1)\) vector of observable explanatory variables and \( x_{2i} \) is an unobservable explanatory variable.

The linear projection of \( x_2 \) on \( x_1 \) is given by \( x_2 = x'_1\delta + v_i \). When \( x_2 \) is unobserved, this implies that the process that determines wages can be expressed as:

\[
w_i = x'_{1i}(\beta_1 + \beta_2\delta) + \beta_2v_i + \epsilon_i \tag{17}
\]

We immediately see that we cannot identify \( \beta_1 \) and \( \beta_2 \delta \) separately, we just can identify a joint parameter \( \tilde{\beta}_1 = \beta_1 + \beta_2 \delta \). This parameter is not a causal or structural, but rather a ‘predictive’ one. For the purpose of finding counterfactual wages and their distribution, this is sufficient.

Secondly, the following set of assumptions on the data-generating process governing wages will be maintained throughout:
Assumption 1.

(1a) **Linearity**: \[ E(y|x_1, x_2) = x'_1\beta_1 + \beta_2 x_2 \]

(1b) **Strict exogeneity**: \[ E(\epsilon_i|x_1, x_2) = 0 \]

(1c) **Full rank**: The matrix \( X = (X_1, x_2) \) has full column rank.

(1d) **Conditional Expectation**: \[ E(x_2|x_1) = x'_1\delta \]

Assumption 2.

**Joint Normality**: \[ \begin{pmatrix} x_2 \\ \epsilon \end{pmatrix} \bigg| x_1 \sim N \left( \begin{pmatrix} x'_1\delta \\ 0 \end{pmatrix}, \begin{pmatrix} \sigma^2_\nu & 0 \\ 0 & \sigma^2_\epsilon \end{pmatrix} \right) \]

Assumption (1b) states that the error term is strictly exogenous to the regressors, conditional on all imaginable relevant explanatory variables, observable or not. Assumption (1d) looks unusual, but ensures that a linear projection of \( x_2 \) on \( x_1 \), written \( x_2 = x'_i\delta_1 + v_i \) actually expresses a conditional mean instead of a mere projection. Assumptions (1a) and (1c) are standard. Assumption 2 implies conditional homoskedasticity of the error term.

Finally, consider also a process that determines the income from self-employment:

\[ \ln \Pi_i = \pi_i = x'_i\tilde{\gamma}_1 + \gamma_2 v_i + \nu_i \quad (18) \]

where \( \tilde{\gamma}_1 \) is immediately written as a reduced-form parameter predicting the relationship between a vector of observables \( x_{1i} \), \( \gamma_2 \) is a coefficient on the unobserved component of income, and \( \nu_i \) is again an i.i.d. error term. Since we cannot separately identify \( \gamma_2 \) and the scale parameter of the distribution of \( v_i \), I normalise \( \gamma_2 = 1.7 \)

In summary, wages and incomes in self-employment are assumed to be given, respectively, by

\[ w_i = x'_i\tilde{\beta}_1 + \beta_2 v_i + \epsilon_i \quad (19) \]

\[ = x'_i\tilde{\beta}_1 + \epsilon_i \quad (20) \]

\[ \pi_i = x'_i\tilde{\gamma}_1 + \gamma_2 v_i + \nu_i \quad (21) \]

\[ = x'_i\tilde{\gamma}_1 + v_i \quad (22) \]

\[ ^7 \text{This implies that } \beta_2 \text{ is also normalised to scale in the following, with some abuse of notation.} \]
3 EMPIRICAL FRAMEWORK

I now discuss, in turn, how counterfactual wages for the self-employed can be identified under different assumptions concerning the correlation structure of the unobservable components of wages and incomes.

3.3.1 OLS wage predictions

If there is conditional independence between unobservables in wage employment and self employment, then \( E(e_i|x_{i1}) = E(e_i|x_{i1}, v_i) \). To justify this assumption, should not be an unobserved factor that determines income in both sectors of employment \((v_i = 0 \forall i)\). Alternatively, this unobserved factor should only be rewarded in one, but not in the other sector. Formally:

**Assumption 3.**

**Conditional independence:** \( v_i \perp e_i | x_{i1} \)

Under this assumption \( \hat{\beta}_1 \) can be identified from a Mincerian wage regression estimated with OLS. The only informative information set is \( \Omega_i = \{x_{i1}\} \).

A linear prediction using this parameter estimate is then an unbiased estimate of the expected wage, given this information set. This is formally stated in

**Proposition 1.** Under Assumptions (1a)-(1d), (3), and i.i.d sampling,

\[
E(\hat{w}_i|x_{i1}) = x'_{i1}E(\hat{\beta}_1|x_{i1}) = x'_{i1}(\beta_1 + \beta_2 \delta_1) = x'_{i1}\beta_1 + E(x_2|x_{i1})\beta_2 = E(w_i|x_{i1})
\]

**Proof.** In Appendix B.

This result enables us the construct the binary measure of outside opportunities (14). Under the stronger assumption of normality, the result about the distribution of the predicted outcome can be obtained:

**Proposition 2.** Under Assumptions (1a)-(1c), (2), and (3)

\[
Pr(w_o > \pi_o|x_{1o}) = \frac{w_o - \hat{w}_o}{\sqrt{\Omega(1 + M_{no})}} \left| X_{1n}, x_{1o} \sim N(0, 1) \right.
\]

with \( M_{no} = x'_{10} \left( \sum_{i=1}^{n} x_{i1}x'_{i1} \right)^{-1} x_{1o} \)

\( \Omega = \sigma^2_\epsilon + \beta_2^2 \sigma^2_v \)
Proof. In Appendix B.

Formally, our counterfactual wage prediction is an out-of-sample prediction (see Casella and Berger 2002). The variance of the prediction for observation \( o \) therefore includes both a variance component due to estimation on a finite sample (captured by the term \( \Omega \cdot M_{no} \)), and the variance of wages \( \Omega \) itself.

3.3.2 Endogenous Switching Model

It is obvious that assumption 3 can easily be violated, and that unobservable components of self-employment income, and wages, can indeed be correlated. Then, selection into wage employment and self-employment is based on comparative advantage (Roy 1951). This implies two things: firstly, the average level of the unobserved component \( x_2 \) (say, ‘ability’) for wage-employed, \( E(x_2|x_1,WE) \), is different from the average level of \( x_2 \) in the population, \( E(x_2|x_1) \).\(^8\) In the notation of the last subsection, \( \delta \) is then different. This means that \( \tilde{\beta}_1 \) is not identified from the subsample of wage-employed only. Secondly, the fact that someone is self-employed is informative about the level of \( v_i \). If this is part of the relevant information set for a counterfactual wage prediction, it should be factored in.

An endogenous switching model (ESM), due to Lee (1978) and Maddala (1983), can be employed to identify the distribution of counterfactual wages under the following assumptions:

**Assumption 4.**

4a Selection mechanism Employment in the wage sector \( (S_i = 1) \) is chosen iff

\[
x_i'\xi + q_i'\zeta + \eta_i > 0
\]

(23)

4b Joint Normality The error terms of (20), (22) and (23) are jointly distributed as

\[
\begin{pmatrix}
v_i \\ e_i \\ \eta_i
\end{pmatrix}
\sim N \left( \begin{pmatrix}
0 \\ 0 \\ 0
\end{pmatrix}, \begin{pmatrix}
\sigma_{ev} & \sigma_{\eta v} \\ \sigma_{ev} & \sigma_{e v} \sigma_{e v} + \sigma_{\eta v} \sigma_{\eta v} \\ \sigma_{\eta v} & \sigma_{\eta v} \\ \sigma_{\eta v} \sigma_{\eta v} & \sigma_{\eta v} \sigma_{\eta v} + \sigma_{\eta v} \sigma_{\eta v}
\end{pmatrix} \right)
\]

\(^8\)In the standard case where selection on unobservables is positive, the average level of \( x_2 \) is higher in the sector with the higher variance. Here, this is self-employment.
4c Information set The relevant information set to predict the counterfactual wage distribution is \( \Psi_i = \{x_{1i}, S_i\} \).

Here, \( q_i \) is a \((R \times 1)\) vector of variables that shift the latent switching variable, but that are excludable from either outcome equation. The symbols \( \beta, \gamma, \xi \) and \( \zeta \) denote parameter vectors of corresponding dimension.

As Lee (1978) and Maddala (1983) explain, the coefficient vectors \( \beta, \gamma \) can be identified and consistently estimated, and so can the variances of \( \sigma_e^2 \) of \( e \) and \( \sigma_v^2 \) of \( v \), and the correlations \( \rho_{ev} \) and \( \rho_{v,v} \). The coefficients \( \zeta \) and \( \xi \) can be identified only up to scale (since the variance of \( v \) is not separately identifiable). The covariance between \( e \) and \( v \) is not identifiable since individuals are only observed in one state at a time.

Furthermore, the model allows us to identify both the unconditional expected wage and the expected potential wage, conditional on being in self-employment. For the latter the unconditional prediction is corrected with the inverse Mills ratio:

\[
E(w_i | S_i = 0) = x_i' \beta - \sigma_e \phi(x_i' \xi + q_i' \zeta) \frac{\phi(x_i' \xi + q_i' \zeta)}{1 - \Phi(x_i' \xi + q_i' \zeta)}
\] (24)

This expression makes clear why assumption (4c) is necessary. We cannot identify \( \rho_{ev} \), and therefore cannot condition wage predictions on an information set that includes an individual’s specific income residual, \( v_i \). We can, however, condition the prediction on the sector an individual actually works in. Expression (24) then states the expected wage, conditional on the information set \((x_{1i}, S_i)\). The model can be estimated by a two-step procedure, or by maximum likelihood, and provides a consistent estimate of this conditional wage, using consistent estimates of population parameters.

3.3.3 Dynamic Search Model (DSM)

We can exploit the panel structure of the data in order to relax the simple assumption on the information set that identification with a static endogenous switching model requires. The dynamic search model (DSM) that I propose models dynamic selection by the transition structure suggested by the theoretical search model. It does not impose a static selection rule. If there are frictions to search, any static composition of employment reflects an unknown mixture between voluntary selection on comparative advantage, and involuntary selection due to frictions.
3 EMPIRICAL FRAMEWORK

The DSM extends the ESM to address some further issues the latter cannot address. Firstly, the switching probability \( \mu \) can be disentangled into two parts: the component of selection based on comparative advantage, that is \( \Pr (w_i > \pi_i | \Psi_i) \), and the component of selection associated with the search friction, \( \lambda \). The endogenous switching model is not able to distinguish between those two components. It rationalises selection into self-employment \( (S_i = 0) \) by a negative shock \( \nu_i \) to selection. On the other hand, \( \rho \eta \nu \) expresses correlation into wage employment, and wages. Therefore, the model might underpredict counterfactual wages for the self-employed. Secondly, the ESM requires a very coarse information set for the conditional wage distribution. By modelling dynamic and not just static selection, the search model allows to identify the parameter \( \rho \eta \nu \), and therefore to include individual-specific residuals \( v_i \) in the information set.

Formally, we observe individual \( i \) in time \( t \in \{1, 2\} \) in sector \( S_{it} \in \{0, 1\} \) having an income:

\[
\begin{align*}
    w_{it} &= x'_{it} \tilde{\beta}_1 + e_{it} \quad \text{iff} \quad S_{it} = 1 \quad (25) \\
    \pi_{it} &= x'_{it} \tilde{\beta}_1 + v_{it} \quad \text{iff} \quad S_{it} = 0 \quad (26)
\end{align*}
\]

I make the following set of assumptions:

Assumption 5.

5a Selection mechanism Transitions between sectors of employment occur according to the following structure:

<table>
<thead>
<tr>
<th>Case</th>
<th>( (S_{11}, S_{12}) )</th>
<th>Transition probability</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>( (0, 1) )</td>
<td>( \lambda \cdot \Pr (w_{i2} &gt; \pi_{i1}) )</td>
</tr>
<tr>
<td>2</td>
<td>( (0, 0) )</td>
<td>( (1 - \lambda) + \lambda \cdot (1 - \Pr (w_{i2} &gt; \pi_{i1})) )</td>
</tr>
<tr>
<td>3</td>
<td>( (1, 1) )</td>
<td>( (1 - \delta) )</td>
</tr>
<tr>
<td>4</td>
<td>( (0, 0) )</td>
<td>( \delta )</td>
</tr>
</tbody>
</table>

5b Joint Normality The error terms of (25), (26) are jointly distributed as

\[
\begin{pmatrix}
    \epsilon_1 \\
    \epsilon_2 \\
    \nu_1 \\
    \nu_2
\end{pmatrix}
\sim N
\begin{pmatrix}
    \sigma^2_\epsilon & \rho \sigma_\epsilon \sigma_\nu & \rho \sigma_\epsilon \cdot \sigma_\nu \\
    \rho \sigma_\epsilon \sigma_\nu & \sigma^2_\nu & \rho \cdot \sigma_\epsilon \cdot \sigma_\nu \\
    \rho \sigma_\epsilon \cdot \sigma_\nu & \rho \cdot \sigma_\epsilon \cdot \sigma_\nu & \sigma^2_\nu \\
    \rho \cdot \sigma_\epsilon \cdot \sigma_\nu & \rho \cdot \sigma_\epsilon \cdot \sigma_\nu & \sigma^2_\nu
\end{pmatrix}
\]
4c Information set 

The relevant information set to predict the counterfactual wage distribution is $\Psi_i = \{x_{1i}, v_i\}$

Note that this transition structure is based on a special case of the ESM where $\xi = 1$, $\zeta = 0$, and $\eta_i = \epsilon_i - \nu_i$. This is the Roy Model, where selection is an immediate function of a comparison of earnings, and hence:

$$\Pr(S_i = 1|\Psi_i) = \Pr(\epsilon_i < x_{2i}^\prime \beta_2 - x_{1i}^\prime \beta_1 + \nu_i|\Psi_i)$$

(27)

In the search model, this probability of drawing an acceptable job offer then is multiplied with the probability of actually receiving a job offer to become $\lambda \cdot \Pr(w_i > \pi_i)$. This selection equation is imposed only on transitions between states of employment, not on static selection into these states within a cross-section.

The model is later estimated using maximum likelihood estimation. Derivation and statement of the likelihood function is relegated to Appendix C.

Finally, I want to place this identification strategy within the literature. Firstly, it relates to a literature on identification of correlated earnings structures. One approach was originally proposed by Lemieux (1998) and recently applied by Suri (2011), Haywood and Falco (2012), and Imbert (2012) in developing country contexts. It identifies correlated, and sector-specific returns to individual unobserved heterogeneity using linear panel methods. That approach does not rely on distributional assumptions, and allows for a general sorting pattern. However, a critical identification assumption is that selection must only occur on time-invariant, and not on transitory, components of unobserved heterogeneity. Since the object of interest in this paper are counterfactual wage distributions, I need to make distributional assumptions. I do not make assumptions on the static sorting pattern, but impose a structure on the dynamic one. In this sense, my approach is more closely related to ‘semi-structural’ approaches of selection correction, following Lee (1978) and Heckman (1979).

Secondly, this identification strategy contrasts with other approaches in the labour literature that estimate the parameters of search frictions. They estimate the job offer arrival rate off the distribution of employment spell durations (e.g. Ridder and van den Berg 2003, Cahuc et al. 2006). Such an approach is infeasible here because the data lack comprehensive information
4 DATA

on employment spells. Rather, I try to identify frictions off the transition rates between employment sectors.

3.4 Identification of effect on outcomes

The ultimate goal of the empirical analysis is to identify the effect of alternative wage employment opportunities on outcomes of the self-employed. The outcomes of interest are transitions out of self-employment into wage-employment, and firm growth. Binary outcomes are regressed on the measures of alternative employment opportunities (29) and (30):

\[
y_{m}^{i,c,t} = \begin{cases} 
1 & \text{if } \alpha_{0} + M_{j}^{i} \alpha_{1}^{j} + \Gamma_{i,c,t}^{t} \alpha_{2} + \tau_{c,t} + \tau_{i,c,t} > 0 \\
0 & \text{if } \alpha_{0} + M_{j}^{i} \alpha_{1}^{j} + \Gamma_{i,c,t}^{t} \alpha_{2} + \tau_{c,t} + \tau_{i,c,t} \leq 0
\end{cases} \tag{28}
\]

where \(y_{m}^{i,c,t}\) denotes the outcome \(m\) for individual \(i\) in city \(c\) and period \(t\), \(M_{j}^{i,c,t}\) the \(j\)th alternative employment opportunities measure, \(\Gamma_{i,c,t}\) a vector of control variables, \(\tau_{c,t}\) city-year dummies, and \(\tau_{i}\) an error term. The main parameter of interest is \(\alpha_{1}^{j}\), the coefficient on the alternative employment opportunity measure \(j\). It can be identified if \(M_{j}^{i,c,t}\) is uncorrelated with the error term \(\tau_{i,c,t}\), conditional on controls. Local labour market conditions (that vary on the city-year level) could drive both income differences (and hence the measures), and outcomes; therefore city-year effects are included. Other idiosyncratic factors could also determine both the measures, and decisions to switch and to grow. However, since the measures are variables constructed from data, they should only directly correlate with the variables that enter them, namely \(x_{1i}\) and \(\pi_{i}\). We can include these variables in the vector of controls, \(\Gamma_{i,c,t}\).

4 Data

4.1 Background

The data used in this analysis consist of a panel survey of urban households in Colombia, the Encuesta Social Longitudinal de Fedesarrollo (Fedesarrollo Social Longitudinal Survey; ESLF). It was collected by the Bogotá-based foundation Fedesarrollo, supported by the Inter-American Development Bank (IADB) and the Bogotá, Bucaramanga and Cali Chambers of Commerce. I use only the 2007-2010 waves because they include a unified module on the labour market. The 2007 wave covers only the Metropolitan
Areas of Bogotá, Bucaramanga and Calí. In 2008, the survey was enlarged to include ten additional cities and Metropolitan Areas.9

The sampling frame consists of the 1993 Census population list of dwellings, households and individuals. Census lists are updated using information from the Colombian National Statistical Office DANE and municipal planning departments. Households are selected into the survey using a thrice-stratified random sampling design: the primary sampling units are Census enumeration areas, secondary sampling units are housing blocks (‘manzanas’), and the final and tertiary sampling units are households. Households attriting from the panel are replaced by new ones.

Generally, for every regression I use the largest sample possible in order to maximise statistical power. For a rough overview, that means pooling cross-sections for OLS estimations, pooling adjacent-year panels for probit and ESM estimations, and using adjacent-year panels (separated by years, and also pooled) for every individual in the DSM estimation. The enlargement of the survey in 2008 may have changed its composition and comparability over time. A similar caveat arises because in the last wave in 2010 fewer households than scheduled could be interviewed in Bucaramanga. The results are robust to all these variations in sampling.

4.2 Descriptive Statistics

The universe of the survey from 2008 onwards comprises of the 13 cities and metropolitan areas constituting the ‘total national urban population’ as defined for statistical purposes by the Colombian National Statistical Office DANE. Table 1 illustrates some general statistics of the urban labour market in Colombia. The labour force participation rate, that is the proportion of the working age population in the labour force, is fairly constant at about 55 percent. A little over 50% of the labour force are wage employed, about a third self-employed, and about 10% are classified as unemployed. Over the years, there is a slight but significant trend of decreasing rates of wage-employment, and increasing rates of unemployment and self-employment.

It should be remembered that those years correspond to the global financial crisis. GDP growth in 2007 was 6.9 percent, then dropped stepwise to 3.5 percent in 2008 and 1.7 percent in 2009 before recovering in 2010.

---

9Barranquilla, Cartagena, Cúcuta, Ibagué, Manizales, Medellín, Montería, Pasto, Pereira, and Villavicencio.
TABLE 1
Composition of the Urban Labour Market in Colombia
2007-2010

<table>
<thead>
<tr>
<th></th>
<th>2007</th>
<th>2008</th>
<th>2009</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>LF Participation Rate</td>
<td>54.4</td>
<td>55.4</td>
<td>54.7</td>
<td>55.4</td>
</tr>
<tr>
<td>% of Labour Force</td>
<td>58.3</td>
<td>54.7</td>
<td>53.0</td>
<td>50.9</td>
</tr>
<tr>
<td>Wage-Employed</td>
<td>31.3</td>
<td>33.7</td>
<td>33.9</td>
<td>35.9</td>
</tr>
<tr>
<td>Self-Employed</td>
<td>8.8</td>
<td>10.0</td>
<td>11.7</td>
<td>12.1</td>
</tr>
</tbody>
</table>

Source: Encuesta Social Longitudinal de Fedesarrollo (ESLF).
Ratios are representative for the total urban national population of Colombia. Wage-employed includes private and government sector employees, domestic and casual workers. Self-Employed includes own-account workers and employers.

with 4 percent and 2011 with 5.9 percent. This adverse macroeconomic environment may explain some of the trends. The global financial crisis constitutes a purely external shock to the Colombian economy. Like other South American economies, it had been on a high growth trajectory for a number of years. The global recession that ensued after the financial crisis in developed countries led to a sharp reduction in GDP growth, with a quick recovery.

I exploit the fact the many the economic downturn brought with it a higher number of involuntary job separations. Labour market transitions across all years are shown in Table 2. ¹⁰ The vertical dimension corresponds to the employment status of an individual in a given year $t$, and the horizontal dimension to the employment status in the following year $t+1$. The table should be read row-wise. Every entry in the main matrix shows the percentage of those in the row category in year $t$ that are in the column category in year $t+1$. For example, 6.6 percent of those out of the labour force switched to wage-employment in the following year. The main diagonal corresponds to inertia: individuals who stay in the same form of employment in two adjacent years.

¹⁰Transition matrices separated by year are available upon request. They show a 20% increase in transitions from wage employment to self-employment in 2008-2009, compared to the previous and following year.
Several stylised facts emerge from the analysis of transitions. First, there is large flexibility in the labour market. Fewer than 75 percent of the wage-employed, and fewer than two thirds of the self-employed remain in their respective occupations. Unemployment is very little persistent, and about half of the unemployed find employment within a year. Second, transitions are asymmetric. The self-employed tend to move either to wage-employment or out of the labour force, but rarely to unemployment. The wage-employed are more likely to switch to self-employment than to any other category. Unemployed are twice as likely to enter wage-employment than self-employment.

In the following, I use only data on the wage-employed and self-employed. In doing so, I abstract from decisions about labour force participation. This approach is justified by the fact that most transitions from wage-employment occur to self-employment, and vice versa. The approach also restricts the analysis to a binary choice situation.

Summary statistics of the variables used in the analysis, and precise variable definitions, can be found in Appendix D. Panel A of Table 11 reports the summary statistics for the explanatory variables. Self-employed individuals are almost a decade older, have more than a year less of formal education, and are more likely to be married as well as head of their household. Interestingly, more than three quarters of both wage-employed and self-employed report being satisfied with their employment, although 41 percent are self-employed by constraint. Panel B describes the outcome variables. A higher
From the static firm size distribution in Figure 1, more than half the firms in the sample (56 percent) are owner-operator firms without other employees. Approximately 21 percent have a single employee, 8 percent have two employees. The ninetieth percentile is at five workers. Less than three percent of the firms in the sample have more than ten workers. This is certainly not a representative sample of Colombian firms. Nevertheless, firms are important for this study insofar as they provide income for, and are affected from decisions by, the self-employed individual who owns the firm.

FIGURE 1
Firm Size Distribution

The transition rates are higher than in Table 2 because here I only consider individuals who are either wage-employed or self-employed in $t$ and $t + 1$. 

\footnote{The transition rates are higher than in Table 2 because here I only consider individuals who are either wage-employed or self-employed in $t$ and $t + 1$.}
5 Reduced-form Results

This section presents the main empirical results obtained with reduced-form methods. Their estimation involves variables constructed from predicted wages. The steps to obtain the predictions are discussed first, followed by outcome regressions and hypotheses tests. I then proceed to several robustness checks. First, I re-estimate wage equations with an endogenous-switching model that tests and corrects for potential selection biases. Secondly, I include firm size and reduce the sample for microentrepreneurs. Thirdly, I look at employment growth in terms of family labour. Fourthly, I check for robustness to using only sub-samples of the data. Fifthly, and finally, I test for credit constraints at the margin of entry into self-employment and for established business owners. The results are preliminary; especially, they were obtained with insufficient repetitions of the computationally expensive bootstrap algorithm (or without bootstrapping at all).

5.1 Predicted Wages

I estimate the Mincer equation (20) by OLS. The logarithm of wages is explained by a function of explanatory variables which is linear in parameters. Coefficients from this estimation are used for wage predictions. Recall that this yields identification under Assumption 4.

I choose a parsimonious representation and regress the log of wages on educational attainment, a second-order polynomial of experience, and on gender. The first two variables are standard in any Mincer equation. Educational attainment is measured in years of schooling and expresses human capital acquired through formal education. Experience expresses human capital obtained through continuous learning on the job. The functional form presumes that returns on education are approximately linear, while the returns profile on experience is allowed to be concave. I add a gender dummy to the human capital variables in order to capture gender-specific differences in wages.

Table 3 presents the results of the Mincer wage regressions. Column (1) includes only the human capital variables. The preferred specification in column (2) adds the gender dummy, which is large in magnitude and highly significant. These coefficients are used for wage predictions. Coefficients are only predictive, and should by no means be interpreted as causal. The sample pools all wage earners from 2007-2010. Since some individuals are
## TABLE 3
**OLS Mincerian Wage Equations**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>OLS</td>
<td>OLS  (ESM sample)</td>
<td>OLS  (Search sample)</td>
<td>OLS</td>
</tr>
<tr>
<td>Years of School</td>
<td>0.100***</td>
<td>0.102***</td>
<td>0.100***</td>
<td>0.097***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.034***</td>
<td>0.033***</td>
<td>0.033***</td>
<td>0.032***</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.002)</td>
<td>(0.003)</td>
</tr>
<tr>
<td>Experience(^2/100)</td>
<td>-0.042***</td>
<td>-0.041***</td>
<td>-0.041***</td>
<td>-0.040***</td>
</tr>
<tr>
<td></td>
<td>(0.004)</td>
<td>(0.003)</td>
<td>(0.004)</td>
<td>(0.005)</td>
</tr>
<tr>
<td>Male</td>
<td>0.245***</td>
<td>0.244***</td>
<td>0.224***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.013)</td>
<td>(0.013)</td>
<td>(0.018)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>11.825***</td>
<td>11.673***</td>
<td>11.689***</td>
<td>11.764***</td>
</tr>
<tr>
<td></td>
<td>(0.032)</td>
<td>(0.033)</td>
<td>(0.033)</td>
<td>(0.050)</td>
</tr>
</tbody>
</table>

|            | 12,880   | 12,880   | 12,063   | 5,916    |
| Adj \(R^2\) | 0.277    | 0.306    | 0.303    | 0.300    |
| Root-MSE   | 0.602    | 0.590    | 0.588    | 0.556    |
| F-stat     | 0.000    | 0.000    | 0.000    | 0.000    |

*** p<0.01 ** p<0.05 * p<0.10. Dependent variable is log of real wages. SE in parentheses. Standard errors clustered at individual level.
included more than once, standard errors are clustered at the individual level. This level of clustering is upheld in all regressions of this section.

Although omitted variables are not a concern for predictions, other sources of endogeneity bias still are. Bias due to reverse causation may still arise. For example, household heads have higher earnings than others, all else being equal. This could be because a household head as the responsible breadwinner needs to work extra hard. Yet, it is equally possible that an individual has been made household head as a result of being the major breadwinner. Then, predictions based on this variable will be inaccurate. Reverse causality bias could also affect other variables potentially endogenous to outcomes, for example marriage status.

The Mincer equation coefficients are robust to sample reductions for the ESM estimation (column 3) and the dynamic estimation in column (4). This justifies using the largest adequate sample for each approach in order to increase efficiency and statistical power.

With the predicted wages I construct the sample counterparts to the measures defined in (13) and (14). This approach is discussed in section (3.2). I define

$$\hat{1}(\pi) = \mathbf{1}(x_i' \hat{\beta}_{1,OLS} > \pi_i)$$

and denote it the ‘binary measure’ and

$$\hat{\Pr}(\pi) = \Phi \left( \frac{x_i' \hat{\beta}_{1,OLS} - \pi_i}{\sqrt{\hat{\Omega}} (1 + M_{ni})} \right),$$

called the ‘continous measure’. The hat symbol serves as a reminder that the variables are sample-dependent estimates, in contrast to (14) and (13), which are population moments.

Table 4 reveals that many self-employed have a good chance of receiving a favourable wage draw. A majority of them can expect to increase their income if they switch to wage-employment.

5.2 Outcome Regressions

With these intermediate results, I can now proceed to testing the hypotheses formulated. I carry out two tests. I first test whether the alternative employ-
ment opportunities measures can predict switching from self-employment to wage-employment. Then, I assess the second hypothesis by testing whether the measures have an effect on firm growth, again, against a no-effect alternative hypothesis. I would interpret such a result as evidence supporting these hypotheses.

In order to implement the tests, the alternative employment opportunities measures (29) and (30) enter as explanatory variables into reduced-form probit models (28). The outcomes are

1. **Switch from SE to WE**: A dummy equal one if a self-employed individual transits to wage-employment between periods \(t\) and \(t+1\), zero if stays self-employed. This variable is restricted to individuals observed in adjacent years in either self-employment or wage-employment.

2. **Grow Firm**: A dummy equal one if self-employed individual adds salaried employees from outside the household to the firm between periods \(t\) and \(t+1\), zero otherwise. This variable is restricted to individuals observed in self-employment in two adjacent years.

Since the outcome regressions include regressors that are constructed using coefficients from first-step regressions, standard errors are obtained with a non-parametric bootstrap procedure.\(^{13}\) I use an algorithm which does

---

\(^{12}\) This is in line with the normality assumptions made so far.

\(^{13}\) Since those regressors are constructed using a finite sample, they introduce a further source of random variation into the regression. Combined with the distribution of the outcome regression error term, it is very likely that a complicated mixture distribution with unknown asymptotic properties arises. In order to still being able to make inference, standard errors are bootstrapped.

---

### Table 4

**OLS Predicted Variables Summary**

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>Mean</th>
<th>SD</th>
<th>I</th>
<th>II</th>
<th>III</th>
</tr>
</thead>
<tbody>
<tr>
<td>(\hat{1}(E(w_i) &gt; \pi_i))</td>
<td>8083</td>
<td>0.652</td>
<td>0.476</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>(\hat{P}(w_i &gt; \pi_i))</td>
<td>8083</td>
<td>0.630</td>
<td>0.311</td>
<td>0.379</td>
<td>0.679</td>
<td>0.931</td>
</tr>
</tbody>
</table>

Sample distribution of predicted measures as defined in equations (29) and (30).
the following upon each bootstrap iteration:\textsuperscript{14}

1. Draw a sample of size $N_O$ with replacement, where $N_O = N_1 + N_2$; that is, the overall sample size $N_O$ is the sum of $N_1$, the size of the OLS subsample of wage-employed and of $N_2$, the size of the probit subsample of self-employed.

2. Estimate a Mincer equation with the wage earner subsample.

3. Predict wages for both subsamples using the coefficients obtained.

4. Construct measures (29) and (30) for the subsample of self-employed.

5. Regress the binary outcome (Switch from SE to WE / Grow Firm) on the measures.

6. Estimate the average marginal effect of the variables of interest.

The basic results are in columns (1)–(4) of Table 5. The binary opportunities measure $\hat{1}(E(w_i) > \pi_i)$ has a positive marginal effect on the switching probability. The relative increase of the likelihood of switching is about 10 percent in magnitude, but not statistically significant. The continuous measure $\hat{Pr}(w_i > \pi_i)$ has a positive marginal effect on switching, which is statistically significant at the ten percent level of significance. Since the explanatory variable is bounded in the unit interval, the marginal effect can be interpreted as: an increase in the measure $\hat{Pr}(w_i > \pi_i)$ by 10 percentage points reduces the switching probability by 0.62 percentage points.\textsuperscript{15} Individuals with a higher probability of receiving an acceptable wage offer are more likely to switch into wage-employment. The test reveals that the constructed measure is informative. However, with a marginal effect very far from one, it is by far not an exhaustive explanation of transitions from self-employment to wage-employment.

The marginal effect of alternative employment opportunities on firm growth is negative and statistically significant. This evidence is consist-

\textsuperscript{14}The procedure upholds the clustering of standard errors at the individual level. This includes using the option \texttt{vce(unconditional)} instead of the delta method for calculating standard errors of marginal effects after the \texttt{margins} command. I need tell the bootstrap that there are two different subsamples, $N_1$ and $N_2$, and define strata accordingly to do so. The algorithm then makes random draws within each stratum, but not across strata.

\textsuperscript{15}In order to avoid confusion, I shall talk of ‘percent’ changes as relative changes in the probability that the outcome takes on the value one, and ‘percentage points’ as absolute increases in this probability.
Table 5: Outcome regressions using OLS-based opportunities measures

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>( \hat{1}(E(w_i) &gt; \pi_i) )</td>
<td>0.020</td>
<td>-0.035**</td>
<td>0.024</td>
<td>-0.038**</td>
<td>0.062*</td>
<td>-0.072*</td>
<td>0.075*</td>
<td>-0.083**</td>
</tr>
<tr>
<td>( \hat{\Pr}(w_i &gt; \pi_i) )</td>
<td>0.062*</td>
<td>-0.072*</td>
<td>0.075*</td>
<td>-0.083**</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

City-Year Dummies

<table>
<thead>
<tr>
<th>N</th>
<th>N</th>
<th>N</th>
<th>N</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
<th>Y</th>
</tr>
</thead>
<tbody>
<tr>
<td>15873</td>
<td>15248</td>
<td>15248</td>
<td>15873</td>
<td>15873</td>
<td>15248</td>
<td>15248</td>
<td></td>
</tr>
<tr>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td></td>
</tr>
<tr>
<td>2,993</td>
<td>2,368</td>
<td>2,368</td>
<td>2,993</td>
<td>2,993</td>
<td>2,368</td>
<td>2,368</td>
<td></td>
</tr>
</tbody>
</table>

\* p<0.10, ** p<0.05, *** p<0.01. Probit Regressions, average marginal effects reported. Dependent variables in column header. All specifications include a constant, specifications (5)-(8) include city and year dummies. SE are obtained by bootstrapping the OLS regression, predictions, and the final probit with 100 replications each. SE are clustered at the individual level.
ent with the combination of both hypotheses: better opportunities in wage-
employment increase the probability of switching, which in turn reduce the
probability of firm growth.

5.3 Robustness

5.3.1 Aggregate Labour Market Shocks

The results could be driven by aggregate shocks to the labour market. For
example, macroeconomic shocks could shift the distribution of earnings. In
this case, individuals that are close to the cutoff might find it worthwhile to
switch upon a shock. Labour market shocks might therefore be the driver
behind our results. This is tested in columns (5)–(8) by including city-year
specific effects. The fixed effects capture aggregate local labour market
shocks. If these shocks were the driving force behind the results, inclusion
them would attenuate the marginal effects of the measures. It would also
reduce their individual significance due to the collinearity of the measures
with the dummies.

A look at the results clearly shows that if anything, the opposite is true:
marginal effects estimates in all specifications gain in absolute magnitude
vis-a-vis the corresponding equation without time and city dummies. They
also ‘gain’ in levels of significance. Hence, I reject the alternative explanation
that aggregate shocks explain the results.

5.3.2 Selection on Unobservables

As discussed in Section 3, selection on unobservables poses a major challenge
for identification of the measures of alternative employment opportunities.
Parameters estimated with OLS only identify the objects of interest under
the set of restrictive assumptions 3. Alternatively, the sets of assumptions
4 and 5 allow for identification using the endogenous switching (ESM) and
dynamic search models (DSM), respectively.

The top panel of table 6 presents the outcome results when counterfactual
wage predictions are based on estimates of the ESM. Results in panel A1,
on the left-hand side, are obtained using maximum-likelihood estimation of
the ESM, and results in panel A2 come from the two-step estimation of
the model. Although there is a lot of instability across estimation methods
in the parameter estimates of the ESM (table available upon request), the
marginal effects of the outcome equation are similar across methods when growth is the outcome. This may be because the parameters are identified from variation in measures across individuals.

**TABLE 6**

<table>
<thead>
<tr>
<th>Unobservables Robustness Regressions</th>
</tr>
</thead>
<tbody>
<tr>
<td>Dependent Variable</td>
</tr>
<tr>
<td></td>
</tr>
<tr>
<td>A1: ESM using MLE</td>
</tr>
<tr>
<td>Pr</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>N2</td>
</tr>
<tr>
<td>B1: DSM 2007/08</td>
</tr>
<tr>
<td>Pr</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>N2</td>
</tr>
<tr>
<td>B2: DSM 2008/09</td>
</tr>
<tr>
<td>Pr</td>
</tr>
<tr>
<td>N</td>
</tr>
<tr>
<td>N2</td>
</tr>
</tbody>
</table>

*** p<0.01 ** p<0.05 * p<0.10. Probit Regressions, average marginal effects. Dependent variables in column header. All specifications include a constant, and city-year (A1, A2, B4) or city (B1, B2, B3) dummies. SE clustered at the individual level. These results are very preliminary, and standard errors have NOT been bootstrapped yet.

Panel B presents the outcome regression results when counterfactual wage distributions are estimated using the DSM. Since pooling ‘cross-sections’ means using the same observation twice here (as a ‘cross-section’ consists of one transition, but income observations from two periods) I also estimate the model separately by year. This results in small sample sizes and large standard errors. Generally, the main results remain robust, though. The marginal example from the pooled regression are even stronger than the OLS results.
There is some variation across years. In 2007/08 and 2009/10, marginal effects are very similar in magnitude to OLS results, although generally not significant. In 2008/09, marginal effects for switching are twice as large, and strongly significant; whereas marginal effects for growth are half the size of baseline results. This is consistent with the large adverse shock in 2009.

Section 6 of the paper discusses the DSM estimation results in more detail.

### 5.3.3 Control regressions

I have found that a higher comparative advantage in wage-employment can explain transitions from self- to wage-employment, as well as firm growth. The results are robust to including a number of controls, and are presented in Table 7.

First, when I control for firm size, the coefficients become larger in absolute magnitude. Firm size itself is significantly related to a higher likelihood of switching, and a lower likelihood of adding employees. This is what one would expect with a concave production function. Furthermore, the results are not consistent with sequential revelation of information (Hopenhayn 1992) as an alternative mechanism.

Second, I also control for income in self-employment. In other words, I compare two individuals that earn the same income in self-employment, but differ in their opportunities to earn income in wage employment. The magnitude of coefficients changes little (for \( \hat{\Pr}(w_i > \pi_i|x_i) \), it slightly increases for switching, and decreases for growth). Higher standard errors render coefficients on \( \hat{1}(E(w_i) > \pi_i|x_i) \) insignificant.

I next control for a number of household variables that attenuate the preference measure coefficients on switching, but render them stronger when firm growth is the outcome. Finally, I confirm that the preference measure does not just pick up direct impacts that the variables used in the wage prediction (education, experience, gender) might have on the outcomes of interest.

Preferences for sectors of employment are not entirely determined by incomes. There could be compensating differentials (Rosen, 1987) which could confound the results. The ESLF survey contains a number of questions on subjective evaluations of employment choices. I construct two subjective
### TABLE 7
#### Control Regressions

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>Switch</th>
<th>Switch</th>
<th>Grow</th>
<th>Grow</th>
<th>Switch</th>
<th>Switch</th>
<th>Grow</th>
<th>Grow</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>SE to WE</td>
<td>SE to WE</td>
<td>Firm</td>
<td>Firm</td>
<td>SE to WE</td>
<td>SE to WE</td>
<td>Firm</td>
<td>Firm</td>
</tr>
<tr>
<td>( \hat{\beta} (1) )</td>
<td>0.026 (0.017)</td>
<td>-0.050*** (0.019)</td>
<td></td>
<td></td>
<td>0.024 (0.017)</td>
<td>0.033 (0.020)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\text{Pr}} (1) )</td>
<td>0.069*** (0.024)</td>
<td>-0.077*** (0.027)</td>
<td></td>
<td></td>
<td>0.079*** (0.025)</td>
<td>-0.070*** (0.030)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Control: Firm Size</td>
<td>0.003*** (0.001)</td>
<td>0.004*** (0.001)</td>
<td>-0.012*** (0.004)</td>
<td>0.013*** (0.001)</td>
<td>0.013*** (0.001)</td>
<td>-0.012*** (0.004)</td>
<td>0.013*** (0.001)</td>
<td>0.013*** (0.001)</td>
</tr>
<tr>
<td>( N )</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
</tr>
<tr>
<td>( N )</td>
<td>2,789</td>
<td>2,789</td>
<td>2,210</td>
<td>2,210</td>
<td>2,701</td>
<td>2,701</td>
<td>2,154</td>
<td>2,154</td>
</tr>
</tbody>
</table>

| \( \hat{\beta} (2) \) | 0.011 (0.017) | -0.041** (0.018) |       |       | 0.027* (0.016) | -0.035** (0.017) |       |       |
| \( \hat{\text{Pr}} (2) \) | 0.036 (0.021) | -0.079** (0.026) |       |       | 0.065*** (0.023) | -0.068*** (0.025) |       |       |
| Control: HH size, head, married, children | 0.003*** (0.001) | 0.004*** (0.001) | -0.012*** (0.004) | 0.013*** (0.001) | 0.013*** (0.001) | -0.012*** (0.004) | 0.013*** (0.001) | 0.013*** (0.001) |
| \( N \) | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 |
| \( N \) | 2,739 | 2,739 | 2,170 | 2,170 | 2,903 | 2,903 | 2,368 | 2,368 |

| \( \hat{\beta} (3) \) | 0.013 (0.017) | -0.027 (0.018) |       |       | 0.0067 (0.016) | -0.027 (0.018) |       |       |
| \( \hat{\text{Pr}} (3) \) | 0.048*** (0.024) | -0.058*** (0.025) |       |       | 0.043* (0.024) | -0.058*** (0.026) |       |       |
| Subjective Preference: Job Dissatisfaction | 0.071*** (0.018) | 0.065*** (0.018) | -0.067*** (0.022) | 0.063*** (0.022) | 0.066*** (0.017) | 0.061*** (0.017) | -0.051*** (0.017) | 0.047*** (0.018) |
| \( N \) | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 | 12,880 |
| \( N \) | 2,500 | 2,500 | 2,366 | 2,366 | 2,919 | 2,919 | 2,309 | 2,309 |

| \( \hat{\beta} (4) \) | 0.021 (0.018) | -0.036* (0.020) |       |       | 0.022 (0.018) | -0.034** (0.017) |       |       |
| \( \hat{\text{Pr}} (4) \) | 0.068*** (0.026) | -0.073*** (0.028) |       |       | 0.068*** (0.025) | -0.072*** (0.023) |       |       |
| Subjective Preference: SE by constraint | 0.071*** (0.018) | 0.065*** (0.018) | -0.067*** (0.022) | 0.063*** (0.022) | 0.066*** (0.017) | 0.061*** (0.017) | -0.051*** (0.017) | 0.047*** (0.018) |
| \( N \) | 10,979 | 10,979 | 10,979 | 10,979 | 10,517 | 10,517 | 10,517 | 10,517 |
| \( N \) | 2,488 | 2,488 | 1,965 | 1,965 | 2,543 | 2,543 | 2,009 | 2,009 |

*** \( p < 0.01 \) ** \( p < 0.05 \) * \( p < 0.10 \). Probit Regressions, average marginal effects. Dependent variables in column header. All specifications include a constant, and city-year dummies. SE obtained by bootstrapping OLS, predictions, and probit with 200 replications each. SE clustered at the individual level. 

32
measures of job satisfaction and subjective sectorial preference, the dummies *Dissatisfied* and *SE by Constraint* (the details are in Appendix D). These measures are included as controls in the outcome probit regressions (28).

The subjective measures, when included as controls, predict outcomes with the expected negative sign: if they increase, switching becomes more likely and growth less likely. The marginal effects of $\hat{\Pr}(w_i > \pi_i|x_i)$ and $\hat{1}(E(w_i) > \pi_i|x_i)$ tend to stay significant and change only very little in magnitude.

The description of the data in section 4.1 lists potential problems for comparability. This affects comparing data from 2007 with later years, and comparing data from Bucaramanga between 2009 and 2010. In order to check that these issues of survey design and sampling do not substantially affect the results, I re-estimate the main regression on subsamples, and results are robust.

Finally, I control for robustness of the results to different forms of wage predictions, in Table 13, Appendix E. I include city and city-year effects in the wage prediction, and I allow for nonlinear wage-schooling profiles by expressing years of schooling in terms of linear spline functions.

5.3.4 Attrition

There is a non-negligible rate of attrition in the survey. About one third of individuals drop out between any given subsequent waves of the survey. Attrition is only weakly correlated with labour market status (SE or WE), and if at all, higher for the wage employed with high earnings. Losing individuals who move into such employment should only lead to a downward bias of my results. On the other hand, attrition is correlated with some of the covariates that enter the regressions. I account for this by weighting observations by the inverse of their predicted attrition probability (Appendix E, Table 14). The change in magnitude of results is similar to the one observed in control regressions.

5.3.5 Normality

In all of this work, I have assumed normality of the wage distribution when estimating the quantiles $\hat{\Pr}(w_i > \pi_i|x_i)$ of the conditional distribution. Overall, this provides a reasonable approximation to the actual distribution of (log) wages. The actual distribution is more skewed to the left, however, and as-
seems a lot of density mass around the minimum wage at approximately the 40th percentile. Here, I assess the sensitivity of my main results with respect to the normality assumption. As an alternative, I express the unconditional wage distribution \( \hat{\Pr}(w_i > \pi_i) \) by the empirical distribution function (EDF) instead of (13) in Table 15, Appendix E. The mean marginal effects of \( \hat{\Pr}(w_i > \pi_i|x_i) \) constructed using either a normal distribution of unconditional wage estimates, or the EDF, are very similar to each other in magnitude. This provides some reassurance that the normality assumption is reasonable. Normality is an essential assumption for the methods in sections ?? and 6 that deal with selection on unobservables.

5.4 Workers from within the household

The definition of firm growth defined growth in terms of adding employees to a firm which are not themselves member of the firm owner’s household. Given that the predictions were obtained from a theoretical framework where investments are (partly) irreversible, this is the right measure to use. Even though firing workers is relatively easy for small employers in the urban informal sector, hiring new workers comes with a certain cost (searching, screening, on-the-job training, etc.) that is especially significant in labour markets with large informational frictions. Conceptually, bearing a hiring cost in all states of the world, but having a benefit from the worker only in some states (when the business is continued) is equivalent to making an irreversible investment that bears positive returns only in some states of the world.

Things look different for workers from within the household, however. Most of the times, they can be ‘hired’ and ‘fired’ for no cost, and they constitute a flexible source of labour to cover spikes in demand, temporary absence of the business owner, etc. In Table 8, I express growth in terms of household workers only. As expected, the marginal effects of both measures on growth of family workers are much smaller, and not significantly different from zero. This is further evidence consistent with the model.

5.5 Credit Constraints

Credit constraints can affect the earnings potential in self-employment. They can also affect the margin of entry from wage-employment into self-employment. Credit constraints at the margin of entry are difficult to reconcile with a waiting queue view of self-employment (Fajnzylber et al. 2006). Furthermore,
my identification strategy as outlined above relies on a sufficient fraction of credit-constrained microentrepreneurs (who are at a firm size below the steady state).

Given these considerations, this section tests for credit constraints. Different tests have been proposed and applied in the literature. Evans and Jovanovic (1989) estimate a structural model for entry into self-employment and test for liquidity constraints with an exclusion restriction on the significance of assets. Banerjee and Duflo (2008) and Karlan and Zinman (2010) exploit the fact that some individuals apply for credit, but are rejected. I employ a simplified version of both approaches. While my regressions lack formal identification, they should give some suggestive evidence whether credit constraints are present.

First, Panel A of Table 9 tests whether assets matter for entry into self-employment. I do not find evidence for this. I run a probit regression

---

**TABLE 8**

**WORKERS FROM WITHIN THE HOUSEHOLD**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch</td>
<td>SE</td>
<td>Switch</td>
<td>SE</td>
<td>Grow</td>
</tr>
<tr>
<td>W</td>
<td></td>
<td></td>
<td></td>
<td>Firm</td>
</tr>
<tr>
<td>( \hat{E}(w_i &gt; \pi_i) )</td>
<td>-0.019</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{Pr}(w_i &gt; \pi_i) )</td>
<td></td>
<td>-0.037</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( N_1 )</td>
<td>12,880</td>
<td>12,880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( N_2 )</td>
<td>2,368</td>
<td>2,368</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** p<0.01 ** p<0.05 * p<0.10. Probit Regressions, average marginal effects. Dependent variables in column header. All specifications include a constant, and city-year dummies. SE obtained by bootstrapping OLS, predictions, and probit with 200 replications each. SE clustered at the individual level.
### TABLE 9
**Credit Constraints**

**A:** Dependent variable is Switch from WE to SE

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef</td>
<td>-0.042***</td>
<td>-0.020</td>
<td>-0.014</td>
<td>-0.065**</td>
<td>-0.011</td>
<td>-0.051***</td>
<td>0.002</td>
<td>0.022***</td>
</tr>
<tr>
<td>Std</td>
<td>(0.010)</td>
<td>(0.016)</td>
<td>(0.025)</td>
<td>(0.024)</td>
<td>(0.015)</td>
<td>(0.012)</td>
<td>(0.027)</td>
<td>(0.009)</td>
</tr>
</tbody>
</table>

N = 5,710

**B:** Dependent variable is Growth.

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Coef</td>
<td>0.015</td>
<td>0.054**</td>
<td>0.148***</td>
<td>0.068**</td>
<td>0.026</td>
<td>0.056***</td>
<td>0.066</td>
<td>-0.016</td>
</tr>
<tr>
<td>Std</td>
<td>(0.018)</td>
<td>(0.026)</td>
<td>(0.045)</td>
<td>(0.039)</td>
<td>(0.025)</td>
<td>(0.025)</td>
<td>(0.045)</td>
<td>(0.045)</td>
</tr>
</tbody>
</table>

N = 2,368

**C:** Dependent variable is Switch from WE to SE, 2007 subsample only

<table>
<thead>
<tr>
<th></th>
<th>(a)</th>
<th>(b)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Applied for Credit</td>
<td>0.042</td>
<td>-0.009</td>
</tr>
<tr>
<td>Std</td>
<td>(0.031)</td>
<td>(0.060)</td>
</tr>
<tr>
<td>Given Credit</td>
<td>0.070</td>
<td>0.068</td>
</tr>
</tbody>
</table>

N = 1,149

*** p<0.01 ** p<0.05 * p<0.10. Probit Regressions, Marginal Effects are reported. All specifications include a constant, year and city dummies, and control for the log income, resp. wage. Standard errors are clustered at the individual level. Regressors are dummies equal one if at least household member uses: (1) a savings account, (2) a current account, (3) another financial savings product (eg fixed-term account), (4) a mortgage, (5) a loan from a non-financial institution, (6) a credit card, (7) participates in a ROSCA, (8) no household member uses any of these financial products.
of entry into self-employment on a number of variables indicating the use of financial products, and controls. If at all, wage-employed individuals who hold financial assets in form of a saving account or who have access to credit in form of a mortgage or a credit card are less likely to enter self-employment. Wage-employed who use no financial product are more likely to do so. Ownership of other household assets such as durable consumption goods cannot predict entry into self-employment.\footnote{Results with non-financial household assets as explanatory variables not reported.}

Use of some financial products does matter for firm growth, though. Holders of current accounts, other savings products, a mortgage or a credit card are more likely to have growing businesses. These results are reported in Panel B of Table 9 and consistent with credit constraints, but may also reflect reverse causality. More successful business owners tend to be more in need of financial products, or qualify more easily for a mortgage.\footnote{In the case of credit cards - most of them issued liberally by retail outlets - the result could also be seen as evidence that better risk-coping capacity is associated with higher growth.}

Panel C of Table 9 shows that individuals who applied for a credit are more likely to switch from wage-employment into self-employment, but only if the application was successful.\footnote{This question was only asked in 2007.} The hazard for individuals who applied but were denied is indistinguishable from the hazard for individuals who did not apply. Moreover, being denied a credit perfectly predicts failure of firm growth, though for a small subsample (N=69). However, credit applications are not evaluated at random, and the results may simply reflect a good job of on part of the credit officers to screen high-able applicants.

In sum, these tests point at little evidence of credit constraints at the margin of entry into self-employment, and stronger evidence of credit constraints for the growth of enterprises. The tests provide supporting evidence in favour of the overall picture. Higher comparative advantage in wage-employment increases the probability of entry into that sector, and reduces the probability of firm growth. Together with little constraints to entry into self-employment, these results are consistent with part of the self-employed forming a waiting queue for wage-employment.
6 Dynamic Search Model

This section presents the results from estimating the dynamic search model from section 3.3.3 using maximum likelihood. In contrast to the previous approaches this model provides an empirical framework to directly estimate and test for search frictions.

For purpose of estimation, more assumptions need to be made about the information structure. Specifically, how should the probability \( \text{Pr}(w_{i,t+1} > \pi_{it}) \) be assessed? In principle, the error term of the first period income equation, \( \nu_{it} = \pi_{it} - E(\pi_{it}|x_{it}) \), is an indicator of unobserved absolute advantage in self-employment. If, like in the Roy model, absolute advantages in self-employment and wage-employment are correlated, employers may be able to recognise and reward it. Only if they do so, then there is a correlation between unobserved components in earnings across sectors. With informational frictions to search present, it is not clear whether this is the case. Fortunately, the alternative scenarios can be tested.

Therefore, I estimate two versions of the dynamic model. In the first version, the probability to receive a favourable job offer depends on the realised absolute advantage in the first period. Individuals with a high absolute advantage in self-employment can also expect a relatively high wage in wage-employment, and vice versa. Note the similarity to the endogenous switching model. Both models correct the expected wage for potential switchers. This correction corresponds to heterogeneity in the unobservable absolute advantage in wage-employment \( \epsilon_i \). In the ESM, the correction term is equal for all individuals within a sector. Its main component is the inverse Mills ratio. In the present model, the dynamic structure imposed allows for individual-specific correction terms based on the realised residual \( \nu_{i1} \) from the first period.

The second version of the model does not condition the expected wage distribution on the realised residual. This version is consistent with employers making wage offers solely based on the information contained in the observable variables \( x_{i2} \). As a result, the earnings residuals across sectors should not be correlated and the constraint \( \rho = 0 \) is imposed in the likelihood function.

The results from estimating the unconstrained and constrained version of the model are given in Table 10. The bottom panel C shows the likelihood
### TABLE 10
**Dynamic Search Model Estimation Results**

#### A - Conditional Model

<table>
<thead>
<tr>
<th></th>
<th>Earnings WE</th>
<th>Earnings SE</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Schooling</td>
<td>0.091***</td>
<td>0.068***</td>
<td>λ</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.035***</td>
<td>0.049***</td>
<td>δ</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.013)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Experience²</td>
<td>-0.000***</td>
<td>-0.001***</td>
<td>ρ</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td>(0.029)</td>
</tr>
<tr>
<td>Male</td>
<td>0.237***</td>
<td>0.529***</td>
<td>ω</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.109)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>11.815***</td>
<td>10.736***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.276)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>σ_ε</th>
<th>σ_ν</th>
<th>ρ_ε</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.572</td>
<td>0.571</td>
<td>0.571</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

|        | 3.115     | 0.079     |
|        | (0.037)   | (0.026)   |

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4771</td>
<td>-19114.5</td>
</tr>
</tbody>
</table>

#### B - Unconditional Model

<table>
<thead>
<tr>
<th></th>
<th>Earnings WE</th>
<th>Earnings SE</th>
<th>Structure</th>
</tr>
</thead>
<tbody>
<tr>
<td>Years of Schooling</td>
<td>0.091***</td>
<td>0.066***</td>
<td>λ</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>Experience</td>
<td>0.035***</td>
<td>0.049***</td>
<td>δ</td>
</tr>
<tr>
<td></td>
<td>(0.002)</td>
<td>(0.013)</td>
<td>(0.006)</td>
</tr>
<tr>
<td>Experience²</td>
<td>-0.000***</td>
<td>-0.001***</td>
<td>ρ</td>
</tr>
<tr>
<td></td>
<td>(0.000)</td>
<td>(0.000)</td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>0.237***</td>
<td>0.551***</td>
<td>ω</td>
</tr>
<tr>
<td></td>
<td>(0.018)</td>
<td>(0.109)</td>
<td></td>
</tr>
<tr>
<td>Constant</td>
<td>11.810***</td>
<td>10.737***</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.043)</td>
<td>(0.278)</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>σ_ε</th>
<th>σ_ν</th>
<th>ρ_ε</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>0.572</td>
<td>0.571</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.006)</td>
<td>(0.014)</td>
<td>(0.014)</td>
</tr>
</tbody>
</table>

|        | 0.079     |
|        | (0.026)   |

<table>
<thead>
<tr>
<th></th>
<th>N</th>
<th>LL</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>4771</td>
<td>-19125.6</td>
</tr>
</tbody>
</table>

#### Likelihood Ratio Test

<table>
<thead>
<tr>
<th>LR Stat</th>
<th>χ²(1)</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>22.27</td>
<td>3.84</td>
</tr>
</tbody>
</table>

*** p<0.01 ** p<0.05 * p<0.10 Maximum Likelihood estimation. Sample reduced to last observed 2-year panel of WE and SE. Standard Errors in Parentheses. Model B is nested in Model A. χ²(1) is the critical value of the χ²-distribution with 1 degree of freedom at the 95% significance level. † Calibrated Parameter. ω is calibrated with the first period sample proportion of self-employed; ρ is calibrated in the restricted (unconditional) model.
test of the two versions of the model. I can confidently reject the hypothesis that the fit of two versions is equal in favour of the unconstrained model in Panel A. In an estimation framework which allows for search frictions, unobservable comparative advantage is still found to matter for selection. The results in Panel A show that the correlation is significantly positive and has a coefficient of 0.139. Frictions are significant and quantitatively important: the probability of a self-employed person to sample a job offer over a year’s time is estimated at 0.288.

The last estimate is robust to imposing the constraint \( \rho = 0 \). Also note that the magnitude of the wage functions coefficients is very similar to the results obtained in OLS and ESM estimation. It is reassuring that the earnings function has been identified in a quite robust manner. My identification strategy relies on robust predictions of wages. Such predictions based on observable characteristics have proven robust across very different model specifications and subsample estimates.

A back-of-the-envelope calculation shows that the simple model is able to explain the stylised facts of transitions remarkably well. I use the estimates of the coefficients of the wage equation, and of the structural parameters, to construct the estimate of \( \Pr (w_{2i} > \pi_{1i} | x_{2i}, \nu_{1i}) \) implied by the model. I then multiply with the estimate for \( \lambda \) to obtain an estimate for \( \mu \), the probability of switching from self-employment to wage employment. The mean of \( \hat{\mu} \) is 0.184, which is very close to the observed proportion of 0.194 of switchers in the dynamic estimation sample.

The estimate of frictions I obtain also appears reasonable given estimates reported in the literature. To my knowledge, this is the first estimate of labour market search frictions for a developing country, and of job search by those currently self-employed. The literature still produces a large variance in results even within OECD countries. Cahuc et al. (2006), for example, find annual rates \( \lambda \) (for search by the unemployed) ranging from 0.05 for unskilled manufacturing to 0.3 for unskilled service workers. Jolivet et al. (2006) report estimates, for unemployed, from 0.33 in Belgium to 0.77 in Germany. They also report estimates for on-the-job-search in wage employment, which are much lower (in the range of 0.02 – 0.10. Ridder and van den Berg (2003), on the other hand, report *monthly* parameters of \( \lambda \) in the range of 0.04 for France to 0.13 for the UK, and even 0.61 for the US.
7 Summary and Outlook

This paper argues that alternative employment opportunities of self-employed affect the performance of microenterprises. In a labour market with search frictions, not all individuals are in their occupation of choice at all times. Individuals differ in their prospects in the wage employment sector, and this affects expectations about firm survival, which in turn influence forward-looking investments.

I formalise this idea with a testable model of investment choice under uncertainty nested into a simple labour market search model. The model predicts that individuals that are more likely to receive attractive job offers (if receiving an offer after all) are more likely to receive self-employment for a wage job. They also have reduced incentives to invest in their firms, and also hire fewer workers.

In the central part of the paper, I derive an empirical strategy to measure the chances in wage-employment among the self-employed, and to test the predictions. I find evidence in favour of both hypotheses. A higher value of the preference measure increases the probability to switch into wage employment, and decreases the probability to grow the firm. This is consistent with the mechanism of the search model.

These results are robust when I correct for aggregate shocks to local labour markets, selection on unobservables, and firm size. Additionally, I devise tests that complement the main results. For outcomes that are very similar, but do not share the features of irreversible investments such as employment of family members, the measures have little explanatory power. Furthermore, I find no evidence for credit constraints at entry into self-employment, which would contradict my findings. I find some evidence of credit constraints for established firms.

I finally construct, derive and estimate a very simple two-period dynamic model of labour market status, transition, and earnings. I find evidence for substantial frictions at the margin of entry to wage-employment, but also for selection based on unobservable comparative advantage.

The validity of most results rests assumptions about the distribution of unobservables, notably joint normality. Identification of the underlying earnings distribution of a multi-sector earnings model where selection is based
REFERENCES

on unobservables almost always relies on such assumptions (Heckman and Honoré 1990). The dynamic model shares these assumptions and extends the endogenous-switching model to a dynamic special case. Still, parameter estimates of the earnings functions are very across all specifications in this paper. This should provide a certain level of confidence in the results which ultimately rely on the identification of counterfactual wages.

All approaches in this paper relied, implicitly or explicitly, on a strong simplification of the discrete dynamic choice model that arises under non-linear, non-additive frictions. The explicit solution of the underlying inter-temporal decision problem, and a structural estimation of a richer model, is the subject of ongoing research. The model of two homogenous sectors could also be enriched by allowing for heterogeneity within sectors. It could also be enlarged to a labour market with three or four sectors (Magnac 1991).

The findings from this research have potential value for policy formulations. This concerns the design of interventions and services for microenterprises on the one hand. On the other hand, policy could ease information flows in the labour market, and ease the degree of search frictions.

References

COLETTO, D., 2010. The informal economy and employment in Brazil:
REFERENCES

*Latin America, modernization, and social changes.* 1 edn. New York: Palgrave Macmillan.


REFERENCES


REFERENCES


Appendix

A Model Derivations

From (4) I obtain:

\[
V(WE, z(k_t)) = u(z(k_t)) + \beta V(WE, z(k_t))
\]

\[
\Leftrightarrow V(SE, k_t) = V(WE, z(k_t)) = \frac{u(z(k_t))}{1 - \beta}
\]

and

\[
V(WE, w) - V(SE, k_t) = \frac{u(w) - u(z(k_t))}{1 - \beta(1 - \delta)}.
\]

With this, I express the SE value function as

\[
V(SE, k_t) = \max_{k_{t+1}} \left\{ u(\pi(k_t) + k_t - k_{t+1}) + \frac{\beta \lambda}{1 - \beta(1 - \delta)} \int_{z(k_{t+1})}^{\infty} \left( u(w) - u(z(k_{t+1})) \right) dF(w) + \beta \frac{u(z(k_{t+1}))}{1 - \beta} \right\}.
\]

I perform integration by parts for a further simplification to

\[
V(SE, k_t) = \max_{k_{t+1}} \left\{ u(\pi(k_t) + k_t - k_{t+1}) + \frac{\beta \lambda}{1 - \beta(1 - \delta)} \int_{z(k_{t+1})}^{\infty} \left( u'(w)(1 - F(w)) \right) dF(w) + \beta \frac{u(z(k_{t+1}))}{1 - \beta} \right\},
\]

from which I can take the first order condition (6) which defines an implicit function in \( k_{t+1} \) and \( \mu = (1 - F(z(k_{t+1}))) = \Pr(w > z(k_{t+1})). \) Its derivatives are

\[
\frac{\partial F}{\partial k_{t+1}} = \beta u''(z(k_{t+1})) \left[ \frac{\partial z(k_{t+1})}{\partial k_{t+1}} \right]^2 \left[ \frac{(1 - \beta) \left[ 1 - \lambda \left( 1 - F(z(k_{t+1})) \right) \right]}{(1 - \beta)(1 - \beta(1 - \delta))} + \beta \delta \right] + \beta u'(z(k_{t+1})) \left[ \frac{\partial^2 z(k_{t+1})}{\partial k_{t+1}^2} \right] \left[ \frac{(1 - \beta) \left[ 1 - \lambda \left( 1 - F(z(k_{t+1})) \right) \right]}{(1 - \beta)(1 - \beta(1 - \delta))} + \beta \delta \right] + \frac{\lambda}{1 - \beta(1 - \delta)} f(z(k_{t+1})) \frac{\partial z(k_{t+1})}{\partial k_{t+1}} + u''(\pi(k_t) + k_t - k_{t+1})
\]

\[
\frac{\partial F}{\partial \mu} = -\frac{\beta \lambda}{1 - \beta(1 - \delta)} \left[ u'(z(k_{t+1})) \frac{\partial z(k_{t+1})}{\partial k_{t+1}} \right].
\]

Noting that \( u'(\cdot) > 0, u''(\cdot) < 0 \) by concavity, \( F(), f() > 0, \beta, \lambda, \delta \in (0, 1) \) and that for intermediate \( k, \frac{\partial z(k_{t+1})}{\partial k_{t+1}} > 0 \) and \( \frac{\partial^2 z(k_{t+1})}{\partial k_{t+1}^2} < 0, \) the sign of both these partial derivatives is negative.
B Proofs of Propositions 1-3

First some general shorthand notation is introduced which shorten the proofs considerably. I have chosen to be consistent with notation and use sum notation throughout, in order to better emphasise where certain assumptions come in; although using only matrix algebra might be more concise.

The OLS estimator of \( w \) regressed on \( x_1 \) only can be written as

\[
\hat{\beta}_{1, \text{OLS}} = (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1}(\sum_{i=1}^{n} x_{ii}w_i)
\]

\[
= (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1}(\sum_{i=1}^{n} x_{ii}(x'_{1i}\hat{\beta}_1 + x_2\beta_2 + \epsilon_i))^{-1}
\]

\[
= \beta_1 + \beta_2\hat{\delta}_1 + \sum_{i=1}^{n} A_i\epsilon_i
\]

and where \( \hat{\delta}_1 \) is the estimator of the coefficient vector \( \delta_1 \) in the linear projection \( x_2 = x'_{1i}\delta_1 + v_i \). This estimator can similarly be written as

\[
\hat{\delta}_1 = (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1}(\sum_{i=1}^{n} x_{ii}x_{2i})
\]

\[
= \delta_1 + \sum_{i=1}^{n} A_i\epsilon_i
\]

which then gives the expression for the first estimator:

\[
\hat{\beta}_{1, \text{OLS}} = \beta_1 + \beta_2\hat{\delta}_1 + \sum_{i=1}^{n} A_i\epsilon_i
\]

with \( c_i = \epsilon_i + \beta_2v_i \)

Denote \( \text{Var}(c_i|\mathbf{X}_1) = \sigma_c^2 \) and \( \text{Var}(v_i|\mathbf{X}_1) = \sigma_v^2 \).

Lemma 1. Under Assumption 1, the following results hold:

a) \( E(c_i|\mathbf{X}_1) = 0 \) and \( E(A_i c_i|\mathbf{X}_1) = A_i E(c_i|\mathbf{X}_1) = 0 \)

b) \( E(c_i^2|\mathbf{X}_1) = \text{Var}(c_i|\mathbf{X}_1) = \sigma_c^2 + \beta_2^2\sigma_v^2 \equiv \Omega \)

c) \( E(c_i c_j|\mathbf{X}_1) = \text{Cov}(c_i|\mathbf{X}_1) = 0 \quad \forall i \neq j \)

d) \( E(\hat{\beta}_{1,\text{OLS}}|\mathbf{X}_1) = \beta_1 + \beta_2\hat{\delta}_1 \)

e) \( \text{Var}(\hat{\beta}_{1,\text{OLS}}|\mathbf{X}_1) = \Omega (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1} \)

Proof.

1. By strict exogeneity, \( E(\epsilon_i + \beta_2v_i|\mathbf{X}_1) = 0 \) and \( E\left( (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1} x_{ii}c_i|\mathbf{X}_1 \right) = (\sum_{i=1}^{n} x_{ii}x'_{1i})^{-1} x_{ii}E(c_i|\mathbf{X}_1) = 0. \)

2. By construction of the linear projection, \( E(\epsilon_i v_i) = 0 \) and then \( \text{Var}(c_i|\mathbf{X}_1) = E(c_i^2|\mathbf{X}_1) - E(c_i|\mathbf{X}_1)^2 = E((\epsilon_i + \beta_2v_i)|\mathbf{X}_1) = \sigma_c^2 + \beta_2^2\sigma_v^2 \equiv \Omega. \)
B.1 Proof of Proposition 1

Consider the prediction \( \hat{w}_i \) of \( w_i \) at a value of \( x_{1i} = x_{10} \). Its expectation obeys

\[
E(\hat{w}_i | X_1, x_{1i} = x_{10}) = E(\hat{w}_i | X_1) = E \left( X'_{10} \hat{\beta}_{1, OLS} | X_1 \right)
\]

appealing to Lemma A1 whereas the expectation of \( w_i \) at of \( x_{1i} = x_{10} \) is

\[
E(w_i | X_1, x_{1i} = x_{10}) = E \left( x'_{10i}\beta_1 + x_{2i}\beta_2 + \epsilon_i | X_1, x_{1i} = x_{10} \right)
\]

\[
= x'_{10} \beta_1 + \beta_2 E(x_{2i} | x_{1i} = x_{10})
\]

\[
= x'_{10} \beta_1 + \beta_2 x_{10} \delta_1
\]

under Assumption 1.

B.2 Proof of Proposition 2

First, I derive the form of the variance of the wage estimate

\[
Var(\hat{w}_i | X_1, x_{1i} = x_{10}) = Var \left( X'_{10} \hat{\beta}_{1, OLS} + \sum A_i \epsilon_i | X_1 \right)
\]

\[
= x'_{10} \left\{ Var \left( \sum A_i \epsilon_i | X_1 \right) \right\} x_{10}
\]

\[
= x'_{10} \left\{ \sum A_i Var(\epsilon_i | X_1) A_i' \right\} x_{10}
\]

\[
= x'_{10} \left\{ \sum x_{1i} x_{1i}' \right\} x_{10}
\]

using all the results from Lemma 1. By Assumption 2, the sum of the normal variables \( \epsilon_i \) and \( \epsilon_i \) is normal, and so is \( \epsilon_i \) which has variance \( \Omega \). Therefore,

\[
\hat{w}_{10} = E(\hat{w}_{10} | X_1) = \frac{\sum A_i \epsilon_i}{\sqrt{x'_{10} \left\{ \sum x_{1i} x_{1i}' \right\} x_{10}}} X_1 \sim N(0, 1).
\]

The variance of an out-of-sample prediction using coefficients estimated on a finite sample comes from two sources. First, 'standard' estimation error; the variance of which is derived about. Secondly, an out-of-sample prediction expresses a
prediction of a random variable which has not yet been realised. The uncertainty surrounding it relates to what Casella and Berger (2002: 557-559) call the difference between a “confidence interval” and a “prediction interval”. This means that for an out-of-sample outcome, we can create an interval in which we can expect to find the outcome, a random variable (Casella and Berger 2002: 559), with a certain probability. It would make no sense within the estimation sample where any observed realisation of the outcome is either inside or outside any fixed interval.

The second source of uncertainty is expressed by the inclusion of \( \epsilon_o \) in the following expression

\[
 w_o - \hat{w}_o = x'_{1o}\beta_1 + \beta_2x_{2o} + \epsilon_o - x'_{1o}(\beta_1 + \beta_2\delta_1) - x'_{1o}\sum_{i=1}^{n} A_i c_i \\
 = \epsilon_o + \beta_2(x_{2o} - x'_{1o}\delta_1) - x'_{1o}\sum_{i=1}^{n} A_i c_i \\
 = \epsilon_o + \beta_2 v_o - \sum_{i=1}^{n} A_i c_i = c_o - x'_{1o}\sum_{i=1}^{n} A_i c_i
\]

the variance of which, noting that the observations indexed \( i \) do not contain observation \( o \) and noting the results in Lemma A1, can be expressed as

\[
 Var (w_o - \hat{w}_o|X_{1n},x_{1o}) = \Omega \left( 1 + x'_{1o} \left\{ \Omega \left( \sum x_{1i}x'_{1i} \right)^{-1} \right\} x_{1o} \right) \\
 = \Omega (1 + M_{no})
\]

C Derivation of Likelihood Function

The likelihood function is derived as follows. First, it should be noted that the likelihood is not separable over time, but it is separable between individuals. The likelihood for every individual, which comprises an observation in each of the two periods, is dependent on the combinations of states, or labour market sectors, the individual is observed in. With 2 periods and 2 states, there are 4 possible cases.

Case 1: SE in period 1, WE in period 2

With \( \pi_{it} = z_{it}\gamma + \nu_{it} \), the density of SE income is \( f(\nu_{i1}) \). All densities are conditional (on \( x_{it} \)), with the conditioning suppressed for easier reading. Similarly, with \( w_{it} = x'_{it}\beta + \epsilon_{it} \), and noting that the earnings observation in the second period is conditional on the error term of the first period, the density of wage in period 2 is \( f(\epsilon_{i2}|\nu_{i1}) \). Together, the product \( f(\nu_{i1}) \cdot f(\epsilon_{i2}|\nu_{i1}) \) is the joint density \( f(\nu_{i1}, \epsilon_{i2}) \) by Bayes’ Rule. The probability mass of observing SE in period 1 is \( (1 - \omega) \) and the probability mass of observing WE in period 2 conditional on SE in period 1 and a residual \( \nu_{i1} \) is \( \lambda \cdot Pr(w_{i2} > \pi_{i1}|\nu_{i1}) \). This implies that the joint likelihood for the individual is \( (1 - \omega) \cdot f(\nu_{i1}) \cdot \lambda \cdot Pr(w_{i2} > \pi_{i1}|\nu_{i1}) \cdot f(\epsilon_{i2}|\nu_{i1}) \). It is manipulated
in the following manner:

\[ L_1 = (1 - \omega) \cdot f (\nu_1) \cdot \lambda \cdot \Pr (w_{i2} > \pi_1 | \nu_1) \cdot f (\epsilon_i | \nu_1) \]

\[ = f (\nu_1) \cdot f (\epsilon_i | \nu_1) \cdot \Pr (x_{i2}^* + \epsilon_i > z_{i1}^* + \nu_1 | \nu_1) \cdot \lambda (1 - \omega) \]

\[ = f (\nu_1, \epsilon_i) \cdot \Pr (-\epsilon_i < x_{i2}^* - z_{i1}^* \gamma + \nu_1 | \nu_1) \cdot \lambda (1 - \omega) \]

\[ = f (\nu_1, \epsilon_i) \cdot \Pr \left( -\epsilon_i + \frac{\sigma_e}{\sigma_e} \nu_1 < x_{i2}^* - z_{i1}^* \gamma + \nu_1 + \frac{\sigma_e}{\sigma_e} \nu_1 \right) \cdot \lambda (1 - \omega) \]

Under the normality assumption (5) the conditional density of \( \epsilon_i | \nu_1 \) is normal with mean \( E (\epsilon_i | \nu_1) = \rho_{\epsilon_2} \nu_1 \) and variance \( Var (\epsilon_i | \nu_1) = \sigma_\epsilon^2 (1 - \rho^2) \). The variable

\[ z = \frac{-\epsilon_i + \frac{\sigma_e}{\sigma_e} \nu_1}{\sigma_e \sqrt{1 - \rho^2}} \sim N (0, 1) \]

then follows a standard normal distribution. Furthermore, the marginal bivariate distribution of \( \nu_1 \) and \( \epsilon_i \) is a bivariate normal distribution. I denote the joint distribution of standardized normal variables \( z_1 = \frac{x_{i1} - \mu_1}{\sigma_1} \) and \( z_2 = \frac{x_{i2} - \mu_2}{\sigma_2} \) with correlation \( \rho \) as

\[ \phi (z_1, z_2; \rho) = \frac{1}{2\pi \rho_1 \rho_2 \sqrt{1 - \rho^2}} \exp \left\{ -\frac{1}{2(1 - \rho^2)} \left[ z_1^2 - 2\rho \cdot z_1 z_2 + z_2^2 \right] \right\}. \]

The likelihood under normality in Case 1 is then

\[ L_1 = \phi \left( \frac{\nu_1}{\sigma_1}, \frac{\epsilon_i}{\sigma_\epsilon}; \rho \right) \cdot \Phi \left( \frac{x_{i2}^* - \pi_1 + \frac{\sigma_e}{\sigma_e} \nu_1}{\sigma_e \sqrt{1 - \rho^2}} \right) \cdot \lambda (1 - \omega) \]

**Case 2: SE in period 1 and 2**

The reasoning follows exactly Case 1. The density of SE income in period 1 is \( f (\nu_1) \) and in period 2 \( f (\nu_2 | \nu_1) \), and the joint density \( f (\nu_1, \nu_2) \). The probability mass of observing SE in period 1 is \( (1 - \omega) \), the probability mass of observing SE in period 2 conditional on SE in period 1 and a residual \( \nu_1 \) is \( 1 - \lambda \cdot Pr (w_{i2} > \pi_1 | \nu_1)\) = \( (1 - \lambda) + \lambda \cdot (1 - Pr (w_{i2} > \pi_1 | \nu_1)) \). Then, the derivation for the likelihood of an individual in Case 2 is

\[ L_2 = (1 - \omega) \cdot f (\nu_1) \cdot [(1 - \lambda) + \lambda \cdot (1 - Pr (w_{i2} > \pi_1 | \nu_1))] \cdot f (\nu_2 | \nu_1) \]

\[ = f (\nu_1, \nu_2) \left[ (1 - \lambda) + \lambda \Pr \left( -\epsilon_i + \frac{\sigma_e}{\sigma_e} \nu_1 < x_{i2}^* - z_{i1}^* \gamma + \nu_1 \right) \cdot \nu_1 \right] (1 - \omega) \]

\[ = \phi \left( \frac{\nu_1}{\sigma_1}, \frac{\nu_2}{\sigma_2}; \rho \right) \left[ (1 - \lambda) + \lambda \cdot \Phi \left( -\frac{x_{i2}^* - \pi_1 + \frac{\sigma_e}{\sigma_e} \nu_1}{\sigma_e \sqrt{1 - \rho^2}} \right) \right] (1 - \omega) \]

noting again the implications of the quadrivariate normality assumption in (5).
Case 3: WE in period 1 and 2

The density of the wage in period 1 is $f(\epsilon_{i1})$ and in period 2 $f(\epsilon_{i2}|\epsilon_{i1})$, jointly $f(\epsilon_{i1}, \epsilon_{i2})$. The probability mass of observing WE in period 1 is $\omega$, the probability mass of observing WE in period 2 conditional on WE in period 1 and a residual $\epsilon_{i1}$ is $(1 - \delta)$. The likelihood of an individual in Case 3 is

$$L^3_i = \omega \cdot f(\epsilon_{i1}) \cdot [1 - \delta] \cdot f(\epsilon_{i2}|\epsilon_{i1})$$
$$= f(\epsilon_{i1}, \epsilon_{i2}) \cdot [1 - \delta] \cdot \omega$$
$$= \phi \left( \frac{\epsilon_{i1}}{\sigma_\epsilon}, \frac{\epsilon_{i2}}{\sigma_\epsilon}; \rho_\epsilon \right) \cdot [1 - \delta] \cdot \omega$$

Case 4: WE in period 1, and SE in period 2

The density of the wage in period 1 is $f(\epsilon_{i1})$ and the density of income in period 2 $f(\nu_{i2}|\epsilon_{i1})$, jointly $f(\epsilon_{i1}, \nu_{i2})$. The probability mass of observing WE in period 1 is $\omega$, the probability mass of observing SE in period 2 conditional on WE in period 1 and a residual $\epsilon_{i1}$ is $\delta$. The likelihood of an individual in Case 4 is

$$L^4_i = \omega \cdot f(\epsilon_{i1}) \cdot \delta \cdot f(\nu_{i2}|\epsilon_{i1})$$
$$= f(\epsilon_{i1}, \nu_{i2}) \cdot \delta \cdot \omega$$
$$= \phi \left( \frac{\epsilon_{i1}}{\sigma_\epsilon}, \frac{\nu_{i2}}{\sigma_\nu}; \rho \right) \cdot \delta \cdot \omega$$

Joint Likelihood

Under assumption of independence across individuals, the joint likelihood is the product of the individual likelihoods:

$$\mathcal{L}(\beta, \gamma, \delta, \lambda, \omega, \sigma_\epsilon, \sigma_\nu, \rho_\epsilon, \rho_\nu) = \prod_{i=1}^{N} \left[ L^1_i \right]^{D^1_i} \left[ L^2_i \right]^{D^2_i} \left[ L^3_i \right]^{D^3_i} \left[ L^4_i \right]^{D^4_i}$$

$D^j_i$ is a dummy variable indicating the case. The log-likelihood is then obtained by taking the natural logarithm of the respective expressions.
### D Summary Statistics and Variable Definitions

**Table 11: Definitions of All Variables Used**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>A: Main Variables</strong></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>Dummy equal one if male, zero if female.</td>
</tr>
<tr>
<td>Age</td>
<td>Age of the individual in years.</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>Years of formal education. This variable was constructed, combining information on the highest level of education attended, and the highest grade completed. For post-secondary and post-graduate education, a standard completion time of 11 and 16 years, respectively, was assumed. For example, an individual who reported 3 years in vocational training was assigned (11 + 3 = 14) years of schooling.</td>
</tr>
<tr>
<td>Experience</td>
<td>Labour market experience in years, constructed as (Age - Years\ of\ Schooling -6).</td>
</tr>
<tr>
<td>HH head</td>
<td>Dummy equal one if household head.</td>
</tr>
<tr>
<td>Married</td>
<td>Dummy equal one for married and cohabitating individuals.</td>
</tr>
<tr>
<td>HH Size</td>
<td>Number of individuals reported to be living in the household.</td>
</tr>
<tr>
<td>Monthly income</td>
<td>Real monthly income (converted to 2010 Colombian pesos using the annual consumer price index) for self-employed individuals; real monthly wage (in 2010 pesos) including overtime and extra pay, but not including fringe benefits (survey question about fringe benefits discontinued in 2007).</td>
</tr>
<tr>
<td>Dissatisfied</td>
<td>Dummy equal one if dissatisfied with current job (answer to yes-or-no question about job satisfaction).</td>
</tr>
<tr>
<td>SE by Constraint</td>
<td>((SE\ only)) Dummy equal zero if a self-employed is so by choice, one if by constraint. Constructed from list of ten options to question about main reason of being self-employed. Top answer for ‘choice’: “better income”; and for ‘constraint’: “It’s the only job I found.”</td>
</tr>
</tbody>
</table>
### B: Savings and Credit Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Savings Account</td>
<td>Dummy equal one if a household member uses the corresponding product. <em>No financial product</em> is only one if no household member uses any of the above.</td>
</tr>
<tr>
<td>Current Account</td>
<td></td>
</tr>
<tr>
<td>Other savings</td>
<td></td>
</tr>
<tr>
<td>Mortgage</td>
<td></td>
</tr>
<tr>
<td>Non-financial credit</td>
<td></td>
</tr>
<tr>
<td>Credit card</td>
<td></td>
</tr>
<tr>
<td>ROSCA</td>
<td></td>
</tr>
<tr>
<td>No financial product</td>
<td></td>
</tr>
<tr>
<td>Applied for Credit</td>
<td>Dummy equal one if a household member applied for a credit in order to install a business, zero if did not apply.</td>
</tr>
<tr>
<td>Denied Credit</td>
<td>Dummy equal one if applied for credit and was denied, zero if did not apply.</td>
</tr>
<tr>
<td>Approved Credit</td>
<td>Dummy equal one if applied for credit and was approved, zero if did not apply.</td>
</tr>
</tbody>
</table>

### C: Outcome Variables

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch WE to SE</td>
<td>(<em>WE only</em>) Dummy equal one if a wage-employed individual transits to self-employment in the next period, zero if stays wage-employed.</td>
</tr>
<tr>
<td>Switch SE to WE</td>
<td>(<em>SE only</em>) Dummy equal one if a self-employed individual transits to wage-employment in the next period, zero if stays wage-employed.</td>
</tr>
<tr>
<td>Grow Firm</td>
<td>(<em>Non-switching SE only</em>) Dummy equal one if self-employed individual adds employees to firm, zero otherwise.</td>
</tr>
</tbody>
</table>

---

53
### TABLE 12

**Summary Statistics of All Variables Used**

<table>
<thead>
<tr>
<th></th>
<th>WE</th>
<th>SE</th>
<th>T-test</th>
<th>p-value</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>N</td>
<td>Mean</td>
<td>Stdev</td>
<td>N</td>
</tr>
<tr>
<td>A</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Male</td>
<td>12,880</td>
<td>0.55</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Age</td>
<td>12,880</td>
<td>35.93</td>
<td>11.84</td>
<td>8,082</td>
</tr>
<tr>
<td>Years of Schooling</td>
<td>12,880</td>
<td>10.74</td>
<td>4.07</td>
<td>8,083</td>
</tr>
<tr>
<td>Experience</td>
<td>12,880</td>
<td>19.19</td>
<td>13.31</td>
<td>8,083</td>
</tr>
<tr>
<td>Household head</td>
<td>12,880</td>
<td>0.36</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Married</td>
<td>12,857</td>
<td>0.48</td>
<td>.</td>
<td>8,075</td>
</tr>
<tr>
<td>HH size</td>
<td>12,063</td>
<td>4.87</td>
<td>1.97</td>
<td>7,545</td>
</tr>
<tr>
<td>Monthly Income †</td>
<td>12,880</td>
<td>803.86</td>
<td>1093.1</td>
<td>8,083</td>
</tr>
<tr>
<td>Monthly Income (log)</td>
<td>12,880</td>
<td>13.32</td>
<td>0.71</td>
<td>8,083</td>
</tr>
<tr>
<td>Dissatisfaction</td>
<td>12,880</td>
<td>0.21</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>SE by Constraint</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>7856</td>
</tr>
<tr>
<td>B</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switch WE to SE</td>
<td>5,178</td>
<td>0.14</td>
<td>.</td>
<td></td>
</tr>
<tr>
<td>Switch SE to WE</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>2,993</td>
</tr>
<tr>
<td>Grow Firm</td>
<td>.</td>
<td>.</td>
<td>.</td>
<td>2,368</td>
</tr>
<tr>
<td>C</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savings Account</td>
<td>12,880</td>
<td>0.47</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Current Account</td>
<td>12,880</td>
<td>0.09</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Other savings</td>
<td>12,880</td>
<td>0.03</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Mortgage</td>
<td>12,880</td>
<td>0.05</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Non-financial credit</td>
<td>12,880</td>
<td>0.09</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Credit card</td>
<td>12,880</td>
<td>0.22</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>ROSCA</td>
<td>12,880</td>
<td>0.03</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>No financial product</td>
<td>12,880</td>
<td>0.41</td>
<td>.</td>
<td>8,083</td>
</tr>
<tr>
<td>Applied for credit</td>
<td>1,901</td>
<td>0.11</td>
<td>.</td>
<td>959</td>
</tr>
<tr>
<td>Was denied credit</td>
<td>1,901</td>
<td>0.03</td>
<td>.</td>
<td>959</td>
</tr>
<tr>
<td>Was given credit</td>
<td>1,901</td>
<td>0.08</td>
<td>.</td>
<td>959</td>
</tr>
</tbody>
</table>

*Source: Encuesta Social Longitudinal de Fedesarrollo (ESLF)*

† In thousands of Colombian Pesos (COP) of 2010. In 2010 the exchange rate was COP 1,898 for one US dollar.

Standard Deviation of binary variables not reported.
### E  Additional Tables of Results

#### TABLE 13  
**ALTERNATIVE WAGE REGRESSION ESTIMATION**

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
<th>(7)</th>
<th>(8)</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Dependent Variable</strong></td>
<td>Switch</td>
<td>Switch</td>
<td>Grow</td>
<td>Grow</td>
<td>Switch</td>
<td>Switch</td>
<td>Grow</td>
<td>Grow</td>
</tr>
<tr>
<td><strong>Variable</strong></td>
<td>Switch</td>
<td>Switch</td>
<td>Firm</td>
<td>Firm</td>
<td>Switch</td>
<td>Switch</td>
<td>Firm</td>
<td>Firm</td>
</tr>
<tr>
<td><strong>Wage Controls: City dummies</strong></td>
<td>0.019 (0.016)</td>
<td>-0.039** (0.015)</td>
<td>0.019 (0.016)</td>
<td>-0.039** (0.017)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pr ()</strong></td>
<td>0.062*** (0.022)</td>
<td>-0.069*** (0.021)</td>
<td>0.065*** (0.022)</td>
<td>-0.073*** (0.025)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_1$</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
</tr>
<tr>
<td>$N_2$</td>
<td>2,789</td>
<td>2,789</td>
<td>2,210</td>
<td>2,210</td>
<td>2,701</td>
<td>2,701</td>
<td>2,154</td>
<td>2,154</td>
</tr>
<tr>
<td><strong>Wage Controls: education spline function</strong></td>
<td>0.020 (0.015)</td>
<td>-0.030* (0.018)</td>
<td>0.025 (0.016)</td>
<td>-0.035** (0.018)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td><strong>Pr ()</strong></td>
<td>0.065*** (0.022)</td>
<td>-0.065*** (0.025)</td>
<td>0.076*** (0.022)</td>
<td>-0.075*** (0.025)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_1$</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
</tr>
<tr>
<td>$N_2$</td>
<td>2,739</td>
<td>2,739</td>
<td>2,170</td>
<td>2,170</td>
<td>2,993</td>
<td>2,993</td>
<td>2,368</td>
<td>2,368</td>
</tr>
</tbody>
</table>

*** p<0.01  ** p<0.05  * p<0.10  
Probit Regressions, average marginal effects. Dependent variables in column header. All specifications include a constant, and city-year dummies. SE obtained by bootstrapping OLS, predictions, and probit with 200 replications each. SE clustered at the individual level.
### TABLE 14

**Inverse Attrition Probability Weights**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Switch SE to WE</td>
<td>0.020</td>
<td>-0.034**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Switch SE to WE</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grow Firm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Grow Firm</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{1}(E(w_i) &gt; \pi_i)$</td>
<td>0.020</td>
<td>-0.034**</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.017)</td>
<td>(0.017)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\hat{Pr}(w_i &gt; \pi_i)$</td>
<td>0.047**</td>
<td>-0.069***</td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.023)</td>
<td>(0.024)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_1$</td>
<td>12,880</td>
<td>12,880</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$N_2$</td>
<td>2,368</td>
<td>2,368</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

*** $p<0.01$ ** $p<0.05$ * $p<0.10$. Probit Regressions, average marginal effects. Weighting with inverse attrition probability weights. Dependent variables in column header. All specifications include a constant, and city-year dummies. SE obtained by bootstrapping OLS, predictions, and probit with 200 replications each. SE clustered at the individual level.
### Table 15

**Robustness: Empirical Distribution Function**

<table>
<thead>
<tr>
<th>Dependent Variable</th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Switch</td>
<td>Switch</td>
<td>Switch</td>
<td>Grow</td>
<td>Grow</td>
<td>Grow</td>
</tr>
<tr>
<td></td>
<td>SE to WE</td>
<td>SE to WE</td>
<td>SE to WE</td>
<td>Firm</td>
<td>Firm</td>
<td>Firm</td>
</tr>
<tr>
<td></td>
<td>Normal</td>
<td>Normal</td>
<td>EDF</td>
<td>Normal</td>
<td>Normal</td>
<td>EDF</td>
</tr>
<tr>
<td>( \hat{1} () )</td>
<td>0.043*** (0.016)</td>
<td></td>
<td></td>
<td>-0.043*** (0.016)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>( \hat{\text{Pr}} () )</td>
<td>0.091*** (0.023)</td>
<td>0.088*** (0.022)</td>
<td></td>
<td>-0.081*** (0.023)</td>
<td>-0.091*** (0.022)</td>
<td></td>
</tr>
<tr>
<td>( N_1 )</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
<td>12,880</td>
</tr>
<tr>
<td>( N_2 )</td>
<td>2,993</td>
<td>2,993</td>
<td>2,993</td>
<td>2,368</td>
<td>2,368</td>
<td>2,368</td>
</tr>
</tbody>
</table>

*** p<0.01 ** p<0.05 * p<0.10. Probit Regressions, average marginal effects. Dependent variables in column header. All specifications include a constant, and city-year dummies. SE obtained by bootstrapping OLS, predictions, and probit with 200 replications each. SE clustered at the individual level.